Journal of Medical Internet Research

Impact Factor (2018): 4.945 - ranked #1 medical informatics journal by Impact Factor
Volume 24 (2022), Issue 2 ISSN: 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Editorial

A Deadly Infodemic: Social Media and the Power of COVID-19 Misinformation (e35552)
Michael Gisondi, Rachel Barber, Jemery Faust, Ali Raja, Matthew Strehlow, Lauren Westafer, Michael Gottlieb. 8

Reviews

Learning Outcomes of Immersive Technologies in Health Care Student Education: Systematic Review of the Literature (e30082)
Grace Ryan, Shauna Callaghan, Anthony Rafferty, Mary Higgins, Eleni Mangina, Fionnuala McAuliffe. 15

Effectiveness of eHealth and mHealth Interventions Supporting Children and Young People Living With Juvenile Idiopathic Arthritis: Systematic Review and Meta-analysis (e30457)
Sonia Butler, Dean Sculley, Derek Santos, Antoni Fellas, Xavier Gironès, Davinder Singh-Grewal, Andrea Coda. 28

Digital Approaches to Automated and Machine Learning Assessments of Hearing: Scoping Review (e32581)
Jan-Willem Wasmann, Leontien Pragt, Robert Eikelboom, De Swanepoel. 52

Implementation of eHealth to Support Assessment and Decision-making for Residents With Dementia in Long-term Care: Systematic Review (e29837)
Juliet Gillam, Nathan Davies, Jesutofunmi Aworinde, Emel Yorganci, Janet Anderson, Catherine Evans. 69

Attrition Within Digital Health Interventions for People With Multiple Sclerosis: Systematic Review and Meta-analysis (e27735)
William Bevens, Tracey Weiland, Kathleen Gray, George Jelinek, Sandra Neate, Steve Simpson-Yap. 87

Digital Health Policy and Programs for Hospital Care in Vietnam: Scoping Review (e32392)
Duc Tran, C Thwaites, Jennifer Van Nuil, Jacob McKnight, An Luu, Chris Paton, Vietnam ICU Translational Applications Laboratory (VITAL). . .
1 0 3

The Effectiveness of Virtual Reality Exposure–Based Cognitive Behavioral Therapy for Severe Anxiety Disorders, Obsessive-Compulsive Disorder, and Posttraumatic Stress Disorder: Meta-analysis (e26736)
Inge van Loenen, Willemijn Scholten, Anna Muntingh, Johannes Smit, Neeltje Batelaan. 134

The Effect of Online Health Information Seeking on Physician-Patient Relationships: Systematic Review (e23354)
Aijing Luo, Lu Qin, Yifeng Yuan, Zhengzijin Yang, Fei Liu, Panhao Huang, Wenzhao Xie. 146

The Effect of Digital Health Interventions on Parents' Mental Health Literacy and Help Seeking for Their Child's Mental Health Problem: Systematic Review (e28771)
Daniel Peyton, Marquelle Goods, Harriet Hiscock. . . . 166

Digital Health Interventions for Weight Management in Children and Adolescents: Systematic Review and Meta-analysis (e30675)
Matina Kouvari, Melina Karipidou, Thomas Tsiampalis, Eirini Mamalaki, Dimitrios Poulimeneas, Eirini Bathrellou, Demosthenes Panagiotakos, Mary Yannakoulia. . . . 181

Posthospitalization Follow-Up of Patients With Heart Failure Using eHealth Solutions: Restricted Systematic Review (e32946)
Ingvild Morken, Marianne Storm, Jon Søreide, Kristin Urstad, Bjørg Karlsen, Anne Husebø. . . . 195

Remotely Delivered Interventions to Support Women With Symptoms of Anxiety in Pregnancy: Mixed Methods Systematic Review and Meta-analysis (e28093)
Kerry Evans, Stefan Rennick-Egglestone, Serena Cox, Yvonne Kuipers, Helen Spiby. . . . 215

Online Health Information Seeking Behaviors Among Older Adults: Systematic Scoping Review (e34790)
Yuxiang Zhao, Mengyuan Zhao, Shijie Song. . . . 234

Sensing Apps and Public Data Sets for Digital Phenotyping of Mental Health: Systematic Review (e28735)
Jean Mendes, Ivan Moura, Pepijn Van de Ven, Davi Viana, Francisco Silva, Luciano Coutinho, Silmar Teixeira, Joel Rodrigues, Ariel Teles. . . 2 5 5

Interactive Visualization Applications in Population Health and Health Services Research: Systematic Scoping Review (e27534)
Jawad Chishtie, Iwona Bielska, Aldo Barrera, Jean-Sebastien Marchand, Muhammad Imran, Syed Tirmizi, Luke Turcotte, Sarah Munce, John Shepherd, Arrani Senthinathan, Monica Cepoiu-Martin, Michael Irvine, Jessica Babineau, Sally Abudiab, Marko Bjelica, Christopher Collins, B Craven, Sara Guilcher, Tara Jeji, Parisa Naraei, Susan Jaglal. . . . 275

Using Social Media to Engage Knowledge Users in Health Research Priority Setting: Scoping Review (e29821)
Surabhi Sivaratnam, Kyobin Hwang, Alyssandra Chee-A-Tow, Lily Ren, Geoffrey Fang, Lindsay Jibb. . . . 297

Digitalized Cognitive Behavioral Interventions for Depressive Symptoms During Pregnancy: Systematic Review (e33337)
Wan Wan Mohd Yunus, Hanna-Maria Matinolli, Otto Waris, Subina Upadhyaya, Miika Vuori, Tarja Korpilahti-Leino, Terja Ristkari, Tarja Koffert, Andre Sourander. . . . 321

Computerized Clinical Decision Support Systems for the Early Detection of Sepsis Among Adult Inpatients: Scoping Review (e31083)
Khalia Ackermann, Jannah Baker, Malcolm Green, Mary Fullick, Hilal Varinli, Johanna Westbrook, Ling Li. . . . 342

Digital Health Technologies to Improve Medication Adherence and Treatment Outcomes in Patients With Tuberculosis: Systematic Review of Randomized Controlled Trials (e33062)
Abdurahman Ridho, Sofa Alfian, Job van Boven, Jutti Levita, Esin Yalcin, Ly Le, Jan-Willem Alffenaar, Eelko Hak, Rizky Abdulah, Ivan Pradipta. 5 3 0

Characteristics of Mobile Health Platforms for Depression and Anxiety: Content Analysis Through a Systematic Review of the Literature and Systematic Search of Two App Stores (e27388)
Qiao Leong, Shreya Sridhar, Agata Blasiak, Xavier Tadeo, GeckHong Yeo, Alexandria Remus, Dean Ho. . . . 607

Viewpoints

Understanding Decision-Making in the Adoption of Digital Health Technology: The Role of Behavioral Economics' Prospect Theory (e32714)
Waqas Khan, Aviv Shachak, Emily Seto. . . . 365

An Ethics Checklist for Digital Health Research in Psychiatry: Viewpoint (e31146)
Francis Shen, Benjamin Silverman, Patrick Monette, Sara Kimble, Scott Rauch, Justin Baker. . . . 369

Innovation Centers in Health Care Delivery Systems: Structures for Success (e33961)
Onil Bhattacharyya, Justin Shapiro, Eric Schneider. . . . 379

Passive Data Use for Ethical Digital Public Health Surveillance in a Postpandemic World (e30524)
John Kilgallon, Ishaan Tewarie, Marike Broekman, Aakanksha Rana, Timothy Smith. . . . 384

Building a Precision Medicine Delivery Platform for Clinics: The University of California, San Francisco, BRIDGE Experience (e34560)
Riley Bove, Erica Schleimer, Paul Sukhanov, Michael Gilson, Sindy Law, Andrew Barnecut, Bruce Miller, Stephen Hauser, Stephan Sanders, Katherine Rankin. . . . 392

Lessons Learned From Beta-Testing a Facebook Group Prototype to Promote Treatment Use in the "Connecting Alaska Native People to Quit Smoking" (CAN Quit) Study (e28704)
Pamela Sinicrope, Colleen Young, Ken Resnicow, Zoe Merritt, Clara McConnell, Christine Hughes, Kathryn Koller, Martha Bock, Paul Decker, Christie Flanagan, Crystal Meade, Timothy Thomas, Judith Prochaska, Christi Patten. . . . 403

The 21st Century Cures Act and Multiuser Electronic Health Record Access: Potential Pitfalls of Information Release (e34085)
Simone Arvisais-Anhalt, May Lau, Christoph Lehmann, A Holmgren, Richard Medford, Charina Ramirez, Clifford Chen. . . . 417

Big Data Health Care Innovations: Performance Dashboarding as a Process of Collective Sensemaking (e30201)
Hilco van Elten, Sandra Sülz, Erik van Raaij, Rik Wehrens. . . . 427

Digital Health Paradox: International Policy Perspectives to Address Increased Health Inequalities for People Living With Disabilities (e33819)
Robin van Kessel, Rok Hrzic, Ella O'Nuallain, Elizabeth Weir, Brian Wong, Michael Anderson, Simon Baron-Cohen, Elias Mossialos. . . . 436

Control Centre for Intensive Care as a Tool for Effective Coordination, Real-Time Monitoring, and Strategic Planning During the COVID-19 Pandemic (e33149)
Martin Komenda, Vladimír erný, Petr Šnajdárek, Mat j Karolyi, Miloš Hejný, Petr Panoška, Ji í Jarkovský, Jakub Gregor, Vojt ch Bulhart, Lenka Šnajdrová, Ond ej Májek, Tomáš Vymazal, Jan Blatný, Ladislav Dušek. . . . 1006

Original Papers

The Effect of Adjunct Telephone Support on Adherence and Outcomes of the Reboot Online Pain Management Program: Randomized Controlled Trial (e30880)
Tania Gardner, Regina Schultz, Hila Haskelberg, Jill Newby, Jane Wheatley, Michael Millard, Steven Faux, Christine Shiner. . . . 449

Effectiveness of Unguided Internet-Based Cognitive Behavioral Therapy and the Three Good Things Exercise for Insomnia: 3-Arm Randomized Controlled Trial (e28747)
Daisuke Sato, Yoichi Sekizawa, Chihiro Sutoh, Yoshiyuki Hirano, Sho Okawa, Motohisa Hirose, Ryo Takemura, Eiji Shimizu. . . . 464

Web-Based Cognitive Testing in Psychiatric Research: Validation and Usability Study (e28233)
Amy Lynham, Ian Jones, James Walters. . . . 481

Internet-Based Audiologist-Guided Cognitive Behavioral Therapy for Tinnitus: Randomized Controlled Trial (e27584)
Eldré W Beukes, Gerhard Andersson, Marc Fagelson, Vinaya Manchaiah. . . . 497

Positive Coping as a Mediator of Mobile Health Intervention Effects on Quality of Life Among People Living With HIV: Secondary Analysis of the Randomized Controlled Trial Run4Love (e25948)
Yu Zeng, Yan Guo, Rainbow Ho, Mengting Zhu, Chengbo Zeng, Aliza Monroe-Wise, Yiran Li, Jiaying Qiao, Hanxi Zhang, Weiping Cai, Linghua Li, Cong Liu. . . . 516

The 3-Month Effectiveness of a Stratified Blended Physiotherapy Intervention in Patients With Nonspecific Low Back Pain: Cluster Randomized Controlled Trial (e31675)
Tjarco Koppenaal, Martijn Pisters, Corelien Kloek, Remco Arensman, Raymond Ostelo, Cindy Veenhof. . . . 543

Long-term Effectiveness of a Smartphone App Combined With a Smart Band on Weight Loss, Physical Activity, and Caloric Intake in a Population With Overweight and Obesity (Evident 3 Study): Randomized Controlled Trial (e30416)
Cristina Lugones-Sanchez, Jose Recio-Rodriguez, Cristina Agudo-Conde, Irene Repiso-Gento, Esther G Adalia, José Ramirez-Manent, Maria Sanchez-Calavera, Emiliano Rodriguez-Sanchez, Manuel Gomez-Marcos, Luis Garcia-Ortiz, EVIDENT 3 Investigators. . . . 559

Blood Pressure Monitoring as a Digital Health Tool for Improving Diabetes Clinical Outcomes: Retrospective Real-world Study (e32923)
Yifat Fundoiano-Hershcovitz, Dror Bacher, Marilyn Ritholz, David Horwitz, Omar Manejwala, Pavel Goldstein. . . . 580

Perceived Acceptability and Experiences of a Digital Psychoeducation and Peer Support Intervention (COPe-support): Interview Study With Carers Supporting Individuals With Psychosis (e27781)
Rachel Batchelor, Sarah Gulshan, Halpana Shritharan, Elen Williams, Claire Henderson, Steve Gillard, Luke Woodham, Victoria Cornelius, Jack Elkes, Jacqueline Sin. . . . 591

Web-Based Research Trends on Child and Adolescent Cancer Survivors Over the Last 5 Years: Text Network Analysis and Topic Modeling Study (e32309)
Hyun-Yong Kim, Kyung-Ah Kang, Suk-Jung Han, Jiyoung Chun. . . . 632

Longitudinal Changes of COVID-19 Symptoms in Social Media: Observational Study (e33959)
Sarah Sarabadani, Gaurav Baruah, Yan Fossat, Jouhyun Jeon. . . . 645

Concerns Around Opposition to the Green Pass in Italy: Social Listening Analysis by Using a Mixed Methods Approach (e34385)
Giovanni Spitale, Nikola Biller-Andorno, Federico Germani. . . . 657

Patients' Perspectives on Transforming Clinical Trial Participation: Large Online Vignette-based Survey (e29691)
Van Nguyen, Philippe Ravaud, Viet Tran, Bridget Young, Isabelle Boutron. . . . 671

Identification of Social Engagement Indicators Associated With Autism Spectrum Disorder Using a Game-Based Mobile App: Comparative Study of Gaze Fixation and Visual Scanning Methods (e31830)
Maya Varma, Peter Washington, Brianna Chrisman, Aaron Kline, Emilie Leblanc, Kelley Paskov, Nate Stockham, Jae-Yoon Jung, Min Sun, Dennis Wall. . . . 684

Effectiveness of Digital Forced-Choice Nudges for Voluntary Data Donation by Health Self-trackers in Germany: Web-Based Experiment (e31363)
Katharina Pilgrim, Sabine Bohnet-Joschko. . . . 696

The Use of Wearable Pulse Oximeters in the Prompt Detection of Hypoxemia and During Movement: Diagnostic Accuracy Study (e28890)
Mauro Santos, Sarah Vollam, Marco Pimentel, Carlos Areia, Louise Young, Cristian Roman, Jody Ede, Philippa Piper, Elizabeth King, Mirae Harford, Akshay Shah, Owen Gustafson, Lionel Tarassenko, Peter Watkinson. 709

Behavioral Efficacy of a Sexual Health Mobile App for Men Who Have Sex With Men: Randomized Controlled Trial of Mobile Messaging for Men (e34574)
Patrick Sullivan, Rob Stephenson, Sabina Hirshfield, Cyra Mehta, Ryan Zahn, Jose Bauermeister, Keith Horvath, Mary Chiasson, Deborah Gelaude, Shelby Mullin, Martin Downing Jr, Evelyn Olansky, Sarah Wiatrek, Erin Rogers, Eli Rosenberg, Aaron Siegler, Gordon Mansergh. . . 7 2 4

Measuring Electronic Health Literacy: Development, Validation, and Test of Measurement Invariance of a Revised German Version of the eHealth Literacy Scale (e28252)
Matthias Marsall, Gerrit Engelmann, Eva-Maria Skoda, Martin Teufel, Alexander Bäuerle. 741

Novel Interactive Tool for Breast and Ovarian Cancer Risk Assessment (Bright Pink Assess Your Risk): Development and Usability Study (e29124)
Elizabeth Hibler, Angela Fought, Kiarri Kershaw, Rebecca Molsberry, Virginia Nowakowski, Deborah Lindner. 757

Health Information Systems for Older Persons in Select Government Tertiary Hospitals and Health Centers in the Philippines: Cross-sectional Study (e29541)
Angely Garcia, Shelley De La Vega, Susan Mercado. 768

Habit and Automaticity in Medical Alert Override: Cohort Study (e23355)
Le Wang, Kim Goh, Adrian Yeow, Hermione Poh, Ke Li, Joannas Yeow, Gamaliel Tan, Christina Soh. 780

SMART COVID Navigator, a Clinical Decision Support Tool for COVID-19 Treatment: Design and Development Study (e29279)
Varun Suraj, Catherine Del Vecchio Fitz, Laura Kleiman, Suresh Bhavnani, Chinmay Jani, Surbhi Shah, Rana McKay, Jeremy Warner, Gil Alterovitz. 792

The Effects of Website Traits and Medical Skepticism on Patients' Willingness to Follow Web-Based Medical Advice: Web-Based Experiment (e29275)
Jennifer Claggett, Brent Kitchens, Maria Paino, Kaitlyn Beisecker Levin. 808

Real-world Implementation of a Smartphone-Based Psychoeducation Program for Bipolar Disorder: Observational Ecological Study (e31565)
Aitana García-Estela, Jordi Cantillo, Natalia Angarita-Osorio, Estanislao Mur-Milà, Gerard Anmella, Víctor Pérez, Eduard Vieta, Diego Hidalgo-Mazzei, Francesc Colom. 824

Prevalence, Factors, and Association of Electronic Communication Use With Patient-Perceived Quality of Care From the 2019 Health Information National Trends Survey 5-Cycle 3: Exploratory Study (e27167)
Rumei Yang, Kai Zeng, Yun Jiang. 843

Examining Diurnal Differences in Multidisciplinary Care Teams at a Pediatric Trauma Center Using Electronic Health Record Data: Social Network Analysis (e30351)
Ashimiyu Durojaiye, James Fackler, Nicolette McGeorge, Kristen Webster, Hadi Kharrazi, Ayse Gurses. 855

Effectiveness, Cost-effectiveness, and Cost-Utility of a Digital Alcohol Moderation Intervention for Cancer Survivors: Health Economic Evaluation and Outcomes of a Pragmatic Randomized Controlled Trial (e30095)
Ajla Mujcic, Matthijs Blankers, Brigitte Boon, Anne Berman, Heleen Riper, Margriet van Laar, Rutger Engels. 870

Implementation of an Electronic Patient-Reported Outcome App for Health-Related Quality of Life in Breast Cancer Patients: Evaluation and Acceptability Analysis in a Two-Center Prospective Trial (e16128)
Joachim Graf, Nina Sickenberger, Katharina Brusniak, Lina Matthies, Thomas Deutsch, Elisabeth Simoes, Claudia Plappert, Lucia Keilmann, Andreas Hartkopf, Christina Walter, Markus Hahn, Tobias Engler, Stephanie Wallwiener, Florian Schuetz, Peter Fasching, Andreas Schneeweiss, Sara Brucker, Markus Wallwiener. 886

Continuous Glucose Monitoring With Low-Carbohydrate Nutritional Coaching to Improve Type 2 Diabetes Control: Randomized Quality Improvement Program (e31184)
Dina Griauzde, Grace Ling, Daniel Wray, Melissa DeJonckheere, Kara Mizokami Stout, Laura Saslow, Jill Fenske, David Serlin, Spring Stonebraker, Tabassum Nisha, Colton Barry, Rodica Pop-Busui, Ananda Sen, Caroline Richardson. 896

Digitally Supported Lifestyle Intervention to Prevent Type 2 Diabetes Through Healthy Habits: Secondary Analysis of Long-Term User Engagement Trajectories in a Randomized Controlled Trial (e31530)
Piia Lavikainen, Elina Mattila, Pilvikki Absetz, Marja Harjumaa, Jaana Lindström, Elina Järvelä-Reijonen, Kirsikka Aittola, Reija Männikkö, Tanja Tilles-Tirkkonen, Niina Lintu, Timo Lakka, Mark van Gils, Jussi Pihlajamäki, Janne Martikainen. 909

Accessibility of Low-cost Insulin From Illegitimate Internet Pharmacies: Cross-sectional Study (e25855)
Benjamin Penley, Lana Minshew, Hui-Han Chen, Stephen Eckel, Sachiko Ozawa. 922

COVID-19 Treatments Sold Online Without Prescription Requirements in the United States: Cross-sectional Study Evaluating Availability, Safety and Marketing of Medications (e27704)
Sachiko Ozawa, Joanna Billings, Yujiao Sun, Sushan Yu, Benjamin Penley. 933

Age- and Sex-Specific Differences in Multimorbidity Patterns and Temporal Trends on Assessing Hospital Discharge Records in Southwest China: Network-Based Study (e27146)
Liya Wang, Hang Qiu, Li Luo, Li Zhou. 946

Acceptability of and Willingness to Take Digital Pills by Patients, the Public, and Health Care Professionals: Qualitative Content Analysis of a Large Online Survey (e25597)
Astrid Chevance, Axel Fortel, Adeline Jouannin, Faustine Denis, Marie-France Mamzer, Philippe Ravaud, Stephanie Sidorkiewicz. 963

The Efficacy of a Smartphone-Based App on Stress Reduction: Randomized Controlled Trial (e28703)
Hyunchan Hwang, Sun Kim, Bo Netterstrøm, Doug Han. 981

The Impact of COVID-19 Confinement on Cognition and Mental Health and Technology Use Among Socially Vulnerable Older People: Retrospective Cohort Study (e30598)
Elena Dura-Perez, Jessica Goodman-Casanova, Amanda Vega-Nuñez, Gloria Guerrero-Pertiñez, Esperanza Varela-Moreno, Maite Garolera, Maria Quintana, Antonio Cuesta-Vargas, Pilar Barnestein-Fonseca, Carlos Gómez Sánchez-Lafuente, Fermin Mayoral-Cleries, Jose Guzman-Parra. 1016

Telehealth Before and During the COVID-19 Pandemic: Analysis of Health Care Workers' Opinions (e29519)
Pascal Nitiema. 1033

Monitoring COVID-19 on Social Media: Development of an End-to-End Natural Language Processing Pipeline Using a Novel Triage and Diagnosis Approach (e30397)
Abul Hasan, Mark Levene, David Weston, Renate Fromson, Nicolas Koslover, Tamara Levene. 1046

COVID-19 Vaccine Tweets After Vaccine Rollout: Sentiment–Based Topic Modeling (e31726)
Luwen Huangfu, Yiwen Mo, Peijie Zhang, Daniel Zeng, Saike He. 1061

A Stanford Conference on Social Media, Ethics, and COVID-19 Misinformation (INFODEMIC): Qualitative Thematic Analysis (e35707)
Michael Gisondi, Daniel Chambers, Tatum La, Alexa Ryan, Adyant Shankar, Athena Xue, Rachel Barber. 1075

Letters to the Editor

Understanding the Impact of Social Media Information and Misinformation Producers on Health Information Seeking. Comment on “Health Information Seeking Behaviors on Social Media During the COVID-19 Pandemic Among American Social Networking Site Users: Survey Study” (e31415)
Hunter Boudreau, Nikhi Singh, Carter Boyd. 997

Digital Inclusion as a Foundation for Health Equity. Comment on “Expanding Video Consultation Services at Pace and Scale in Scotland During the COVID-19 Pandemic: National Mixed Methods Case Study” (e34702)
Christopher Weatherburn. 999

Authors’ Reply: Understanding the Impact of Social Media Information and Misinformation Producers on Health Information Seeking. Comment on “Health Information Seeking Behaviors on Social Media During the COVID-19 Pandemic Among American Social Networking Site Users: Survey Study” (e31569)
Stephen Neely, Christina Eldredge, Ronald Sanders. 1001

Getting a Vaccine, Jab, or Vax Is More Than a Regular Expression. Comment on “COVID-19 Vaccine-Related Discussion on Twitter: Topic Modeling and Sentiment Analysis” (e31978)
Jack Cummins. 1004

Editorial

A Deadly Infodemic: Social Media and the Power of COVID-19 Misinformation

Michael A Gisondi[1], MD; Rachel Barber[2]; Jemery Samuel Faust[3], MS, MD; Ali Raja[4], MBA, MD; Matthew C Strehlow[5], MD; Lauren M Westafer[6], DO; Michael Gottlieb[7], MD

[1]The Precision Education and Assessment Research Lab, Department of Emergency Medicine, Stanford University, Palo Alto, CA, United States
[2]Stanford University, Stanford, CA, United States
[3]Department of Emergency Medicine, Brigham and Womens Hospital, Harvard University, Boston, MA, United States
[4]Department of Emergency Medicine, Massachusetts General Hospital, Harvard University, Boston, MA, United States
[5]Stanford Emergency Medicine International, Department of Emergency Medicine, Stanford University, Stanford, CA, United States
[6]Department of Emergency Medicine, University of Massachusetts, Worchester, MA, United States
[7]Department of Emergency Medicine, Rush University, Chicago, IL, United States

Corresponding Author:
Michael A Gisondi, MD
The Precision Education and Assessment Research Lab
Department of Emergency Medicine
Stanford University
900 Welch Road - Suite 350
Palo Alto, CA, 94304
United States
Phone: 1 650 721 4023
Email: mgisondi@stanford.edu

Abstract

COVID-19 is currently the third leading cause of death in the United States, and unvaccinated people continue to die in high numbers. Vaccine hesitancy and vaccine refusal are fueled by COVID-19 misinformation and disinformation on social media platforms. This online *COVID-19 infodemic* has deadly consequences. In this editorial, the authors examine the roles that social media companies play in the COVID-19 infodemic and their obligations to end it. They describe how *fake news* about the virus developed on social media and acknowledge the initially muted response by the scientific community to counteract misinformation. The authors then challenge social media companies to better mitigate the COVID-19 infodemic, describing legal and ethical imperatives to do so. They close with recommendations for better partnerships with community influencers and implementation scientists, and they provide the next steps for all readers to consider. This guest editorial accompanies the Journal of Medical Internet Research special theme issue, “Social Media, Ethics, and COVID-19 Misinformation.”

(J Med Internet Res 2022;24(2):e35552) doi:10.2196/35552

KEYWORDS

COVID-19; social media; misinformation; disinformation; infodemic; ethics; vaccination; vaccine hesitancy; infoveillance; vaccine

The COVID-19 Infodemic

The COVID-19 pandemic continues to have a substantial impact worldwide, with over 266 million diagnosed cases and over 5 million deaths [1]. In 2021, depending on the month, COVID-19 was either the first, second, or third leading cause of death in the United States, alongside heart disease and cancer [2]. People are still dying from COVID-19 despite a vaccine surplus in wealthy countries, public health interventions to curb viral transmission, new therapeutic options, and the heroic efforts of frontline care providers. Why? Although we initially focused on a deadly and contagious virus, we were simultaneously overwhelmed by the deadly and contagious impact of online misinformation and disinformation about that virus [3]. Much like the COVID-19 pandemic itself, we face a widespread disease with long-term consequences: the COVID-19 infodemic.

The World Health Organization defines an infodemic as “too much information or false and misleading information” that “causes confusion, risk taking behaviors...and mistrust of health

officials" [4]. The United Nations Educational, Scientific, and Cultural Organization considers *fake news* a general term for false information that can be further defined by intentionality [5]. Misinformation consists of "information that is false but not created with the intention of causing harm," whereas disinformation is "information that is false and deliberately created to harm a person, social group, organisation, or country" often orchestrated with financial or political motives [5]. Both are prevalent across all social media platforms [6]. Together, these serve to undermine trust in governmental interventions, public health responses, expert guidance, and scientific facts about COVID-19 [7,8]. Accordingly, we define the COVID-19 infodemic as the overwhelming amount of complex and often contradictory information available about COVID-19, inclusive of substantial fake news about the origins of the virus, treatment options unsupported by rigorous clinical data, and baseless claims regarding adverse effects of lifesaving vaccines; these false narratives may be spread by authoritative institutions or influencers who are otherwise thought to be trustworthy, and they play a substantial role in shaping views and influencing human behaviors that can lead to poor health outcomes.

The clinical impact of the COVID-19 infodemic is profound. Effective strategies such as masking and social distancing have been undermined to the detriment of those at greatest risk. With several effective vaccines now available for SARS-CoV-2, vaccine hesitancy and vaccine refusal—two distinct problems with different causes and different solutions—have become major issues. Vaccine hesitancy is prolonged deliberation or delay in accepting vaccination, even when supply is ample; this differs from vaccine refusal, which is defined by the specific intent not to vaccinate, similar to the "anti-vax" movement adherents, in which people refuse all vaccines including childhood vaccinations. Both vaccine hesitancy and refusal are fueled by misinformation on social media, and vaccine misinformation that initially manifests offline can quickly spread to social media platforms; the misinformation exchange is bidirectional [9]. In fact, the US Surgeon General warned in 2021 that misinformation is the greatest threat to COVID-19 vaccination efforts [10]. COVID-19 misinformation and disinformation on social media increases vaccine hesitancy, lowers vaccination rates, and causes preventable deaths, especially among certain demographic populations [11,12]. The COVID-19 infodemic remains deadly, and we must act.

To address this, the Stanford University Ethics, Society, and Technology Hub and the Stanford Department of Emergency Medicine cosponsored INFODEMIC: A Stanford Conference on Social Media and COVID-19 Misinformation. INFODEMIC convened experts from the fields of social media, medicine, public health, and biomedical ethics with a goal of identifying new best practices to combat COVID-19 misinformation online [13]. The corresponding Journal of Medical Information Research theme issue, "Social Media, Ethics, and COVID-19 Misinformation" builds upon this work to discuss the impact of this infodemic and approaches to ending it. In this editorial, we will examine the role of social media companies (executives, financiers, leaders, and users) in health misinformation and their obligations to mitigate the COVID-19 infodemic.

The Role of Social Media

Recognition of social media's power to propagate fake health news came well before COVID-19, notably surrounding topics such as tobacco use, vaping, and recreational drug use [6]. However, 2020 was the *year of online disinformation*, with political and scientific misinformation and disinformation often reinforcing one another [14]. Social media companies and platform users both played a substantial role in the birth of the COVID-19 infodemic that year. The internet propagates knowledge rapidly and globally, typically without checks for accuracy, and facilitates the current infodemic. Social media companies attempt to self-police erroneous content on their platforms with variable success, both due to the overwhelming amount of COVID-19 information they must process and their reluctance to censor their users' posts. Information filters swiftly through various avenues on the internet, often accessed via search engines and social media algorithms. Google is the dominant search engine with over 3.5 billion global searches each day, allowing individuals to retrieve information from a wide array of sources [15]. Although it may appear that users pull information, search engines prioritize certain results, in effect pushing relevant information to the user [16]. Social media algorithms push selected content to billions of users as well (Table 1) [17]. The proprietary algorithms used by social media companies are routinely exploited to spread COVID-19 misinformation and disinformation, with certain content repeatedly presented to users who have specific profiling characteristics or search histories. These algorithms could be better optimized to reduce the online trafficking of harmful information that risks the public health.

Table 1. Approximate numbers of monthly users of several social media platforms.

Social media platform (company, location)	Approximate number of users
Facebook (Meta Platforms, Inc; Menlo Park, California)	3 billion
YouTube (Google LLC; San Bruno, California)	2.3 billion
WhatsApp (Meta Platforms, Inc; Menlo Park, California)	2 billion
Instagram (Meta Platforms, Inc; Menlo Park, California)	1.4 billion
Twitter (Twitter; San Francisco, California)	400 million

In addition to social networking, an increasing number of users consume news on social media platforms compared with traditional media outlets [18]. Individuals engage in social circles and networks on these platforms virtually, and they do

not leave—their beliefs are reinforced by information chosen for them and others like them by computer algorithms. Without intervention, there is exposure to new content but little to no exposure to new knowledge or ways of thinking. The algorithms used by social media companies create news echo chambers that can serve as a vector for misinformation by amplifying low credibility information sources. During the early COVID-19 pandemic, low credibility sources dominated both Twitter and Facebook posts related to COVID-19, topping traditional news and media outlets [19]. Online *bots* further confuse users and reduce their ability to discern truth from fake news. Bots are computer codes designed to appear as user profiles or credible news sources but are instead weapons for disinformation. Social media companies struggle to identify and remove bots that use even the simplest artificial intelligence, which take advantage of platform data and social media push algorithms. Thus, it is unsurprising that social media platforms fuel hoaxes and misinformation about the etiology and origins of COVID-19, its treatment, and its prevention through vaccines [19]. Social media companies could invest greater resources to combat these agents of the infodemic.

Moreover, health misinformation is not confined to COVID-19. In a 2021 systematic review, the greatest prevalence of health misinformation was found on Twitter and related to smoking, drugs, and vaccines [6]. Who is to blame? Many fingers can be pointed, and social media companies are among the culpable. *Top-down* misinformation from celebrities and other public figures that are allowed on these platforms exacerbates the problem. Celebrities account for 20% of online misinformation and 70% of the attention of platform users, compared to noncelebrity posts [19]. Social media companies benefit from increased user activity, and celebrity influencers are engaging. These attention-grabbing individuals enjoy unfettered reach to users because social media companies rarely place limits on their messaging, even when that messaging includes erroneous facts about COVID-19. The blurred lines between factual news and entertainment and falsehoods about COVID-19 could be labeled for users by social media companies through better oversight of their platforms.

An Obligation to Act

US-based social media companies are legally regulated by the US Federal Trade Commission in the same ways that any other US-based businesses are regulated. However, they are not subject to federal *social media regulations* of any kind—because there are none [20]. Social media platforms are private companies who set their own internal regulatory policies and are not subject to oversight by the US Federal Communications Commission (FCC) nor, specifically, Section 230 of the Communications Act of 1934 [21]. Subsection (c) (1) of the Act maintains that social media companies do not act as publishers of information, as do other media entities, absolving them of an obligation to monitor user-generated content on their platforms. There are no legal mandates to the manner or methods by which they self-govern, and any actions by these companies are simply made in good faith. However, they do maintain internal policies, some of which are intended to curb dissemination of different types of information that are harmful to the public welfare, ranging from COVID-19 misinformation to communications among terrorist organizations. These Good Samaritan policies determine what constitutes an acceptable use of their platforms and draw a line limiting certain types of social expression [22].

That said, social media companies are mostly unregulated, and some claim they should remain so, rather than be subject to FCC oversight as are television and radio companies. Current actions by these companies are voluntary, not compulsory, and often in response to external pressures. However, if companies do not meaningfully address misinformation and disinformation on their platforms, government oversight should be considered. Internal policies are an important first step, yet we continue to see blatantly false information that contradicts scientific evidence regularly posted across most social media platforms [3]. For example, a content analysis study that evaluated 1225 fake news stories found that social media platforms were responsible for disseminating half (50.5%) of the identifiable misinformation [3]. Social media companies are partly responsible for fueling the COVID-19 infodemic, and we believe that ongoing inaction or inadequate action to address it keeps them complicit. Given the stakes, failure to address health misinformation and disinformation should be viewed as a public health crisis, and the commensurate response should include government oversight of social media companies (similar to other media sources) in the name of the public good.

Beyond law and public policy, there are other interventions to consider. We maintain that there is an ethical obligation for social media companies to act. Bioethicists recognize the broader public health consequences of social media use and how bias is determined by the specific design and implementation of social media platforms [23]. Ethical frameworks guide moral decision-making and action/inaction, two of which are especially relevant for social media companies [24]. First, utilitarian ethics calls for decisions that positively affect the greatest number of people. This is also a bedrock of public health. The application of utilitarian ethics suggests that companies should make socially conscious decisions, even when inconvenient [25]. For example, social media companies can and should redesign their algorithms that have propagated the infodemic, even if such changes risk ad revenue and profits. Similarly, censorship of celebrity users who disseminate misinformation might decrease user engagement and activity with a platform, but these actions are needed to address the infodemic. Second, virtue ethics is reasoning based on the virtues of what makes a good person, or in this case a good company. Kaptein [26] defined corporate virtue ethics that include clarity, transparency, and sanctionability, among others. A *good* social media company that earnestly engages in self-regulation would exhibit many other virtues including honesty, courage, self-control, and integrity. When considering corporate virtues, and what makes a *good* social media company, it is worth noting corporate vices in the absence of *good*: deficiency, ambiguity, subversiveness, and opaqueness [26]. These vices are commonly associated with the current practices of social media companies, specifically when considering their algorithms that repeatedly drive dangerous content to users. These algorithms reinforce

COVID-19 misinformation for some users and cloister them from reports on legitimate scientific evidence.

It does not have to be this way. There is hope for "good social media companies," ones which take bold actions in the face of a global pandemic, and perhaps at their own expense. We think that social media companies could become allies after all, especially as their power to enact social change is unprecedented. Indeed, many companies are working to mitigate misinformation already. For instance, Twitter, Facebook, and Google Search remove content, add warning labels, and deactivate misleading accounts that promote disinformation. However, these companies must navigate the substantial influx of constantly changing information about COVID-19, struggling to discern between deceptive content, scientifically inaccurate *alternative facts*, and even genuine scientific disagreement [27]. That struggle deserves consideration. For example, a claim that SARS-CoV-2 is airborne may have been flagged as misinformation or disinformation in January 2020. Today, in early 2022, most scientists believe that the virus is airborne at least in certain environments and conditions. Would a social media company that removed a post claiming that SARS-CoV-2 is airborne have harmed the public or prevented important academic debate in the early days of the pandemic?

Acknowledging the genuine complexity of the situation does not undermine but rather illuminates the urgency of our call to action. In fact, recognizing the difficulties that social media companies have in fairly adjudicating misinformation implies that far more financial and human resources will need to be marshaled to sufficiently address the problem. Wealthy companies have invested some capitol toward these efforts, but they are capable of so much more. The volume of misinformation is impressive, and removed information is often quickly replaced by similarly harmful messages within a platform, stifling progress. In essence, social media companies are treading water, at best, and additional resources and initiatives are warranted. Such efforts may need to be strategically directed toward specific aspects of the infodemic, such as misinformation about the clinical severity of the virus or disinformation about the efficacy of the vaccine. Different strategies are warranted to address a range of fixed beliefs that may have developed at different stages of the pandemic; a singular approach by a social media company may be insufficient to change minds about the origins of the virus, its transmissibility, and its prevention, all at once. Each type of misinformation deserves a unique message in response, and these messages must be tailored to the cultural differences of users in certain communities.

Addressing the infodemic does not fall solely to social media companies, and we cannot rely only on a few people or a small number of entities to battle it. Key potential change agents also include elected government representatives, public health officials, research scientists, journalists, clinical ethicists, and physicians. However, for many constituencies, a general distrust of government and an underappreciation of science is magnified in part by social media, requiring a truly multidisciplinary response to the challenge [28]. Therefore, we believe that major impact could also result from effective messaging delivered by trusted community leaders who can, for example, reach communities of vaccine hesitant individuals online and in person: ministers, youth counselors, teachers, social and mental health workers, frontline workers, and others.

Finally, implementation scientists have been inadequately used as resources. Implementation science is the study of methods used to introduce research findings into a health care context. With respect to COVID-19, the methods used to introduce a new vaccine to the entire world's population represented an important and frequently missed opportunity [29]. During 2020, significant attention was given to the rapid development and testing of COVID-19 vaccines; relatively less attention was given to equally crucial areas, including how to equitably manufacture vaccines to scale and how to equitably distribute them [30]. The COVID-19 vaccine is widely accessible throughout most industrialized nations now, but the potential influence of implementation science remains no less important. A key tenet of implementation science is the use of different strategies to target different patient populations [31]. Especially in the face of stiffening vaccine hesitancy in certain communities and subpopulations, we believe that implementation scientists should be engaged to help guide strategies and actions of social media companies, and similarly to help other change agents such as elected officials. There is a social imperative and an ethical imperative to embrace the best available evidence and methods needed to improve COVID-19 vaccination rates. We recommend that social media companies seek the advisement of experts in implementation science as they develop strategies to combat vaccine and other COVID-19 misinformation.

Next Steps

Health misinformation and disinformation have been increasing rapidly for over a decade [14]. During the COVID-19 pandemic, substantial attention was initially focused on ensuring the distribution of the vaccines themselves but, unfortunately, *not the distribution of reliable information nor the mitigation of harmful misinformation and disinformation.* This has had long-lasting effects. Over time, some who have said "I won't wear a mask" now say, "I won't take a vaccine." Going forward, it is imperative that we move beyond our roles as scientists in a laboratory, physicians on the frontline, or strategists at social media companies. We must seek to expand our influence in health education and public health messaging more broadly. We must exit the silo of the house of medicine and meet patients and the public where they are at: online. As we do this, we need to rely on sound strategies gleaned from education and leadership literature to reach our patients effectively [8,32]. We need to identify evidence-based interventions that effectively dismantle online misinformation and then implement them [14].

If we want to create meaningful change, we cannot merely rely on the progress of clinical science alone. We must consider how best to implement and disseminate new discoveries to the public via social networks and offline communities. For many in science and medicine, this may mean engaging with mass media for the first time. That means personalizing our direct outreach to patients and communities, engaging with empathy, and seeking to understand before seeking change [33]. Moreover, we must resist a paternalistic approach in which we protect

patients from information but rather empower them to seek reliable information and make informed choices about their health. Fortunately, there are frameworks that can guide us. Bautista et al [33] proposed a two-step conceptual model for physicians seeking to refute misinformation—step 1: identification of the types and sources of health misinformation; and step 2: attempting to make private and public corrections, done strategically and respectfully. Meanwhile, Chou et al [34] urged physicians to partner with social media companies and influencers to address health misinformation online, teach the public to recognize potential misinformation, and cultivate better trust toward the medical community. Finally, Walter et al [35] confirmed that interventions to correct health misinformation are the most successful when they come from experts in a given field.

Complicated problems call for collaborative approaches. Social media companies, medical professionals, researchers, implementation scientists, and trusted messengers must form synergistic partnerships to successfully combat the COVID-19 infodemic and health misinformation and disinformation generally. Rather than focusing on assigning blame, change agents can be most effective by demonstrating their willingness to act and implement new best practices, regardless of whether or not they previously contributed to some of the problems we face today.

As you read the articles in this special theme issue of the Journal of Medical Internet Research, we urge you to reflect on expanding your own contributions beyond your current working environment. In Textbox 1, we offer several actionable next steps for social media companies and health care providers to combat COVID-19 misinformation. Consider how we all can better address the current COVID-19 infodemic and combat and prevent future ones. To truly win this battle, we must urgently convert our expertise into the right words and the right actions. Whether we find our patients in the clinical environment or on social media, we must protect them from the harms of misinformation and disinformation and help them benefit from the lifesaving medical and health information that we have to offer.

Textbox 1. Actions to address the COVID-19 infodemic.

Recommendations for social media companies

- Redesign social media algorithms to reduce the spread of COVID-19 misinformation
- Identify and remove harmful bots from platforms
- Censor sources of COVID-19 misinformation and disinformation
- Label erroneous content
- Promote sound science
- Support public health efforts
- Target culturally appropriate messaging to specific communities
- Direct users to local health clinics and COVID-19 resources

Recommendations for health care providers

- Engage patients on social media
- Offer COVID-19 content expertise to social media companies and online news media
- Commit to posting public health messaging online
- Identify and implement evidence-based interventions to combat health misinformation
- Partner with online influencers to disseminate accurate COVID-19 information
- Provide expert advice to mass media outlets
- Personalize direct outreach to patients and communities
- Seek to understand patients with empathy before seeking behavior change
- Empower patients to seek reliable health information and make informed choices
- Create synergistic partnerships with leaders in other disciplines

Authors' Contributions

MAG, RB, JSF, AR, MCS, LMW, and MG contributed to this editorial. All but RB are guest section editors for the Journal of Medical Internet Research special theme issue, “Social Media, Ethics, and COVID-19 Misinformation.”

Conflicts of Interest

None declared.

References

1. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). Johns Hopkins Coronavirus Resource Center. URL: https://coronavirus.jhu.edu/map.html [accessed 2021-12-07]
2. Average number of daily deaths in the United States among the leading causes of death from March 2020 to November 2021. Statista. URL: https://www.statista.com/statistics/1254560/leading-causes-of-death-in-the-us-average-number-daily/ [accessed 2021-11-13]
3. Naeem SB, Bhatti R, Khan A. An exploration of how fake news is taking over social media and putting public health at risk. Health Info Libr J 2021 Jun;38(2):143-149 [FREE Full text] [doi: 10.1111/hir.12320] [Medline: 32657000]
4. Infodemic. World Health Organization. URL: https://www.who.int/health-topics/infodemic [accessed 2021-11-13]
5. Journalism, 'Fake News' and disinformation: a handbook for journalism education and training. UNESCO. 2019. URL: https://en.unesco.org/fightfakenews [accessed 2021-11-14]
6. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. J Med Internet Res 2021 Jan 20;23(1):e17187 [FREE Full text] [doi: 10.2196/17187] [Medline: 33470931]
7. Gottlieb M, Dyer S. Information and disinformation: social media in the COVID-19 crisis. Acad Emerg Med 2020 Jul;27(7):640-641. [doi: 10.1111/acem.14036] [Medline: 32474977]
8. Sheng AY, Gottlieb M, Welsh L. Leveraging learner-centered educational frameworks to combat health mis/disinformation. AEM Educ Train 2021 Aug;5(4):e10711. [doi: 10.1002/aet2.10711] [Medline: 34926972]
9. Pullan S, Dey M. Vaccine hesitancy and anti-vaccination in the time of COVID-19: a Google Trends analysis. Vaccine 2021 Apr 01;39(14):1877-1881 [FREE Full text] [doi: 10.1016/j.vaccine.2021.03.019] [Medline: 33715904]
10. Holcombe M. Surgeon General warns misinformation is the greatest threat to COVID-19 vaccination efforts. CNN Philippines. URL: https://www.cnn.ph/world/2021/6/26/surgeon-general-warns-misinformation-threat-to-COVID-19-vaccination.html [accessed 2021-10-26]
11. de Albuquerque Veloso Machado M, Roberts B, Wong BLH, van Kessel R, Mossialos E. The relationship between the COVID-19 pandemic and vaccine hesitancy: a scoping review of literature until August 2021. Front Public Health 2021;9:747787. [doi: 10.3389/fpubh.2021.747787] [Medline: 34650953]
12. Razai MS, Osama T, McKechnie DGJ, Majeed A. Covid-19 vaccine hesitancy among ethnic minority groups. BMJ 2021 Feb 26;372:n513. [doi: 10.1136/bmj.n513] [Medline: 33637577]
13. Stanford Infodemic Conference. 2021. URL: https://www.stanfordinfodemic.org/ [accessed 2021-11-13]
14. Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. Soc Sci Med 2019 Nov;240:112552 [FREE Full text] [doi: 10.1016/j.socscimed.2019.112552] [Medline: 31561111]
15. Google Search statistics. Internet Live Stats. URL: https://www.internetlivestats.com/google-search-statistics/ [accessed 2021-11-03]
16. Makhortykh M, Urman A, Ulloa R. How search engines disseminate information about COVID-19 and why they should do better. HKS Misinformation Review. 2020 May 11. URL: https://misinforeview.hks.harvard.edu/article/how-search-engines-disseminate-information-about-covid-19-and-why-they-should-do-better/ [accessed 2021-11-03]
17. Dean B. Backlinko. 2021 Oct 10. URL: https://backlinko.com/social-media-users [accessed 2021-11-08]
18. Gottfried J, Shearer E. News use across social media platforms 2016. Pew Research Center. 2020 Aug 27. URL: https://www.pewresearch.org/journalism/2016/05/26/news-use-across-social-media-platforms-2016/ [accessed 2021-11-21]
19. Brennan J, Simon F, Howard PN, Nielsen RS. Types, sources, and claims of COVID-19 misinformation. Reuters Institute for the Study of Journalism. 2020 Apr 07. URL: https://reutersinstitute.politics.ox.ac.uk/types-sources-and-claims-covid-19-misinformation [accessed 2021-11-28]
20. Coldewey D. Who regulates social media? TechCrunch. 2020 Oct 19. URL: https://techcrunch.com/2020/10/19/who-regulates-social-media/ [accessed 2021-10-28]
21. Johnson Jr TM. The FCC's Authority to Interpret Section 230 of the Communications Act. Federal Communications Commission. 2020 Oct. URL: https://www.fcc.gov/news-events/blog/2020/10/21/fccs-authority-interpret-section-230-communications-act [accessed 2021-10-28]
22. Hintz A. Social media censorship, privitized regulation, and new restrictions to protest and dissent. In: Dencik L, Leistert O, editors. Critical Perspectives on Social Media and Protest: Between Control and Emancipation. Washington, DC: Rowman & Littlefield; 2015.
23. Terrasse M, Gorin M, Sisti D. Social media, e-health, and medical ethics. Hastings Cent Rep 2019 Jan;49(1):24-33. [doi: 10.1002/hast.975] [Medline: 30790306]
24. Ethical theories. The Arthur W Page Center: Public Relations Ethics Training. URL: https://pagecentertraining.psu.edu/public-relations-ethics/introduction-to-public-relations-ethics/lesson-1/ethical-theories/ [accessed 2021-11-13]

25. Gustafson A. In defense of a utilitarian business ethic. Business Soc Rev 2013 Sep 03;118(3):325-360. [doi: 10.1111/basr.12013]
26. Kaptein M. When organizations are too good: applying Aristotle's doctrine of the mean to the corporate ethical virtues model. Business Ethics Eur Rev 2017 Feb 22;26(3):300-311. [doi: 10.1111/beer.12147]
27. Conway: Press Secretary gave 'Alternative Facts'. NBC News. 2017 Jan 22. URL: https://www.nbcnews.com/meet-the-press/video/conway-press-secretary-gave-alternative-facts-860142147643 [accessed 2021-11-27]
28. Huber B, Barnidge M, Gil de Zúñiga H, Liu J. Fostering public trust in science: the role of social media. Public Underst Sci 2019 Oct;28(7):759-777 [FREE Full text] [doi: 10.1177/0963662519869097] [Medline: 31524092]
29. Wensing M. Implementation science in healthcare: introduction and perspective. Z Evid Fortbild Qual Gesundhwes 2015;109(2):97-102. [doi: 10.1016/j.zefq.2015.02.014] [Medline: 26028446]
30. Wensing M, Sales A, Armstrong R, Wilson P. Implementation science in times of Covid-19. Implement Sci 2020 Jun 08;15(1):42 [FREE Full text] [doi: 10.1186/s13012-020-01006-x] [Medline: 32513292]
31. Bauer MS, Kirchner J. Implementation science: what is it and why should I care? Psychiatry Res 2020 Jan;283:112376 [FREE Full text] [doi: 10.1016/j.psychres.2019.04.025] [Medline: 31036287]
32. Kotter J. Leading change: why transformation efforts fail. Harvard Business Rev 2007;85(1):2-12.
33. Bautista JR, Zhang Y, Gwizdka J. Healthcare professionals' acts of correcting health misinformation on social media. Int J Med Inform 2021 Apr;148:104375. [doi: 10.1016/j.ijmedinf.2021.104375] [Medline: 33461008]
34. Chou WS, Oh A, Klein WMP. Addressing health-related misinformation on social media. JAMA 2018 Dec 18;320(23):2417-2418. [doi: 10.1001/jama.2018.16865] [Medline: 30428002]
35. Walter N, Brooks JJ, Saucier CJ, Suresh S. Evaluating the impact of attempts to correct health misinformation on social media: a meta-analysis. Health Commun 2021 Nov;36(13):1776-1784. [doi: 10.1080/10410236.2020.1794553] [Medline: 32762260]

Abbreviations

FCC: Federal Communications Commission

Edited by A Mavragani; submitted 08.12.21; peer-reviewed by A Sheng, A Merritt, H Ayatollahi; comments to author 31.12.21; revised version received 03.01.22; accepted 06.01.22; published 01.02.22.

Please cite as:
Gisondi MA, Barber R, Faust JS, Raja A, Strehlow MC, Westafer LM, Gottlieb M
A Deadly Infodemic: Social Media and the Power of COVID-19 Misinformation
J Med Internet Res 2022;24(2):e35552
URL: https://www.jmir.org/2022/2/e35552
doi:10.2196/35552
PMID:35007204

Review

Learning Outcomes of Immersive Technologies in Health Care Student Education: Systematic Review of the Literature

Grace V Ryan[1], MUDR, MRCPI; Shauna Callaghan[1], RM, PhD; Anthony Rafferty[1], BSc, MBChB, PhD; Mary F Higgins[1], MSc, MD, MRCPI, MRCOG; Eleni Mangina[2], HDip, MSc, PhD; Fionnuala McAuliffe[1], MD, MRCPI, MRCOG

[1]Perinatal Research Centre, Obstetrics and Gynaecology, School of Medicine, University College Dublin, Dublin, Ireland

[2]School of Computer Science, University College Dublin, Dublin, Ireland

Corresponding Author:
Fionnuala McAuliffe, MD, MRCPI, MRCOG
Perinatal Research Centre
Obstetrics and Gynaecology
School of Medicine, University College Dublin
The National Maternity Hospital
65/66 Lower Mount Street Dublin
Dublin, D02 NX40
Ireland
Phone: 353 1 6373216
Email: fionnuala.mcauliffe@ucd.ie

Abstract

Background: There is a lack of evidence in the literature regarding the learning outcomes of immersive technologies as educational tools for teaching university-level health care students.

Objective: The aim of this review is to assess the learning outcomes of immersive technologies compared with traditional learning modalities with regard to knowledge and the participants' learning experience in medical, midwifery, and nursing preclinical university education.

Methods: A systematic review was conducted according to the Cochrane Collaboration guidelines. Randomized controlled trials comparing traditional learning methods with virtual, augmented, or mixed reality for the education of medicine, nursing, or midwifery students were evaluated. The identified studies were screened by 2 authors independently. Disagreements were discussed with a third reviewer. The quality of evidence was assessed using the Medical Education Research Study Quality Instrument (MERSQI). The review protocol was registered with PROSPERO (International Prospective Register of Systematic Reviews) in April 2020.

Results: Of 15,627 studies, 29 (0.19%) randomized controlled trials (N=2722 students) were included and evaluated using the MERSQI tool. Knowledge gain was found to be equal when immersive technologies were compared with traditional learning modalities; however, the learning experience increased with immersive technologies. The mean MERSQI score was 12.64 (SD 1.6), the median was 12.50, and the mode was 13.50. Immersive technology was predominantly used to teach clinical skills (15/29, 52%), and virtual reality (22/29, 76%) was the most commonly used form of immersive technology. Knowledge was the primary outcome in 97% (28/29) of studies. Approximately 66% (19/29) of studies used validated instruments and scales to assess secondary learning outcomes, including satisfaction, self-efficacy, engagement, and perceptions of the learning experience. Of the 29 studies, 19 (66%) included medical students (1706/2722, 62.67%), 8 (28%) included nursing students (727/2722, 26.71%), and 2 (7%) included both medical and nursing students (289/2722, 10.62%). There were no studies involving midwifery students. The studies were based on the following disciplines: anatomy, basic clinical skills and history-taking skills, neurology, respiratory medicine, acute medicine, dermatology, communication skills, internal medicine, and emergency medicine.

Conclusions: Virtual, augmented, and mixed reality play an important role in the education of preclinical medical and nursing university students. When compared with traditional educational modalities, the learning gain is equal with immersive technologies. Learning outcomes such as student satisfaction, self-efficacy, and engagement all increase with the use of immersive technology, suggesting that it is an optimal tool for education.

(J Med Internet Res 2022;24(2):e30082) doi:10.2196/30082

KEYWORDS

Virtual Reality; Augmented Reality; Mixed Reality; Learning Outcomes; Medical Education; Nursing Education; Midwifery Education; Systematic Review

Introduction

Background

Educational technology is changing the way in which we learn today, and its purpose is to ultimately improve education [1,2]. The addition of educational technology to a curriculum needs to be developed and guided by informed, evidence-based research. Educational technology includes instructional software such as virtual reality (VR), augmented reality (AR), and mixed reality (MR), known collectively as immersive technology [3]. Immersive technologies should be built around effective teaching methods that provide an appropriate learning method and learning outcome [4]. Immersive technologies are thought to provide pedagogy based on the constructivist theory and experiential learning, creating an environment that aids visual learners and enables students to learn by doing, develop creativity, and increase understanding of invisible concepts [5]. The Association for Medical Education in Europe has previously published guidance on e-learning in medical education: "Designs for effective medical e-learning, therefore, need to mirror the dynamics and details of real-world practice as well as affording effective learning opportunities" [6].

Immersive technologies are defined as devices that provide sensory stimuli to provide a sense of realism and immersion in the interactions with the computer-generated world [7]. VR is a technology that allows the user to explore and manipulate computer-generated real or artificial 3D multimedia sensory environments in real time. It dates back to 1956, when Morton Heilig, a cinematographer, developed *the Sensorama*, a display box first used for background scenes in the Hollywood motion picture industry. This was the first head-mounted display to be developed. In the mid-1960s, Ivan Sutherland, an American Computer Scientist, went on to develop the concepts of VR further. He described *The Ultimate Display*, a VR system that could simulate reality [8], and his paper described core concepts that are the foundation of VR use today. Owing to the heterogeneity of the terminology used to describe VR, we can characterize VR as a "collection of hardware such as Personal Computer (PC), HMDs and tracking sensors, as well as software to deliver an immersive experience" [9]. In comparison, AR is an interactive experience of a real-world environment where the objects that reside in the real world are *augmented* by computer-generated perceptual information. Historically, the development of AR started in the 1960s; however, the term *AR* was not established until 1990. Although VR and AR share many technical aspects, the main difference is that AR does not construct a fully artificial environment and simply overlays computer-generated images onto images of the real world. Therefore, it uses machines that allow a physical view of the surrounding environment to be visible but enhanced with virtual images [10]. Finally, MR is the merging of real and virtual worlds to produce new environments and visualizations where physical and digital objects coexist and interact in real time [11].

Objective

To date, there has been a multitude of publications detailing the development and implementation of immersive tools, in addition to demonstrating the benefits of VR, AR and MR technology in medical, nursing and midwifery education [12-17]. This technology is thought to provide increased engagement and understanding during learning coupled with feedback mechanisms and design capabilities of varying difficulty levels [5]. In addition, it facilitates practice without the risk of human harm and also helps build professional skills and teamwork [18,19]. However, there is a paucity of evidence on the learning outcomes of these innovative educational tools.

As outlined by the Digital Health Education Collaboration, there is a need for a strong evidence base to guide the development of immersive educational tools so that the learning goals and outcomes are in line with national and international standards [20]. There have been an increasing number of systematic reviews documenting the use, application, and effectiveness of VR, AR, and MR in an effort to establish an evidence-based network of research for use in medical education. However, the results have been mixed; including a systematic review that looked at the effectiveness of AR in medical education which found that there was insufficient evidence to recommend its implementation into the curriculum. Similarly, another review looked at serious games used in medical education and found that the evidence was moderate to support the use of immersive technology, stating that it should not replace traditional learning tools [21-25]. Immersive technologies are used mainly as educational tools for complex topics such as anatomy and embryology and are thought to enhance the learning experience [26,27]. Are VR, AR, and MR as effective in delivering knowledge as well as an enhanced learning experience in comparison with traditional teaching tools such as 2D didactic presentations?

Therefore, the aim of this systematic review is to assess the learning outcomes of VR, AR, and MR across 3 health care student disciplines—medicine, nursing, and midwifery education—compared with traditional learning modalities. The learning outcomes include knowledge, skill development, and the learning perceptions of students, including satisfaction and self-confidence in learning along with engagement and motivational factors.

Methods

Purpose and Protocol

A systematic review of the available scientific literature was conducted to assess the learning outcomes associated with the application of VR, AR, and MR as educational tools compared

with traditional learning modalities for medical, nursing, and midwifery students in preclinical university education. The review protocol was registered with PROSPERO (International Prospective Register of Systematic Reviews) in April 2020 (CRD42020154598). The search results were reported according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [28] and the Cochrane Collaboration guidelines [29].

Eligibility Criteria

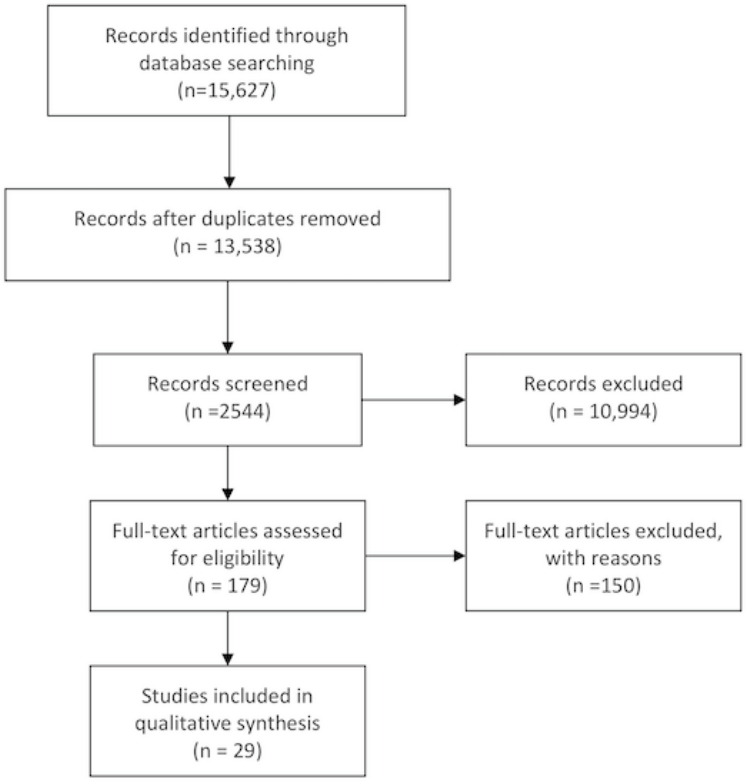
The eligibility criteria were based on the Population, Intervention, Comparison, and Outcomes criteria. The population selected for this review included preclinical university students enrolled in three educational disciplines: medicine, nursing, or midwifery courses only. Randomized controlled trials (RCTs) that implemented VR, AR, or MR technology in comparison with a control method were included. Owing to the heterogeneity of the definitions surrounding VR applications, we restricted inclusion to interactive 3D models requiring a headset, virtual patients (VPs), or VR learning environments. The primary outcomes included knowledge and reference to the learning experience, which involved engagement, satisfaction, and perceived learning experience. Only studies published in English were included.

Search Strategy

A large-scale search was undertaken because of the wide use of the various terminology to describe VR, AR, and MR and the technology surrounding their use in health care student education. The following method was used to identify empirical studies for inclusion in the systematic review. We conducted a comprehensive computerized database search of full-text articles published in English. Only RCTs assessing learning outcomes using VR, AR, or MR technologies in comparison with traditional learning models were included. The reason for this was that we wanted to review the learning outcomes of immersive technologies compared with traditional learning outcomes, including knowledge and learner experience. The fundamental study design of an RCT requires a control and an intervention group; therefore, we selected these types of studies for this review. Searches were conducted with predefined search terms (Multimedia Appendix 1) using the following electronic databases: PubMed, Embase, Web of Science, CINAHL, and ERIC. Medical Subject Headings terms included *virtual reality*, *augmented reality*, *educational technology*, *imaging*, *three dimensional*, *education*, and *teaching materials*. Search terms were connected using Boolean operators *AND* and *OR* to capture all relevant article suggestions. The latest search was conducted on March 8, 2021.

Databases were downloaded to EndNote reference manager software (Clarivate Analytics), which recorded citations and identified duplicates. A spreadsheet was created to record decisions and comments. Screening of articles was conducted by 2 researchers (GR and SC) in an unblinded, standardized approach (independently and in parallel). Titles and abstracts of studies sourced from electronic databases were reviewed according to the inclusion and exclusion criteria previously described. Subsequently, full texts of the included articles from the initial screening process were reviewed for eligibility according to the described inclusion and exclusion criteria. Differences of opinion were resolved through conversations between the reviewers. A schematic stepwise algorithm for the search strategy is shown in Figure 1.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram adapted for this study.



Data Collection

Microsoft Excel was used to build a data extraction form, which was divided into three categories: (1) study identification, (2) analysis of learning outcomes, and (3) study design. The first section included bibliographic information, the country of origin of the study, and a demographic description of the participants. The second section examined learning outcomes related to teaching strategies, relationships between technologies, and learning objectives. The third section evaluated the methodological quality of the study design.

Study Quality Assessment

We used the Medical Education Research Study Quality Instrument (MERSQI) to evaluate the study design of the RCTs [30]. The MERSQI is divided into several domains, including evaluation of study design, sampling, data type, validity, data analysis, and outcomes. The learning outcomes are based on the hierarchy of educational outcomes by Kirkpatrick and Kirkpatrick [31], which adopts a constructional framework using a 4-level model for evaluating educational effectiveness. The first level describes the participants' perception of the learning experience; knowledge, skills, and attitudinal change are assessed in the second level; changes in behavior are evaluated in the third level; and changes in health care or patient outcomes are evaluated in the fourth level.

Data Analysis

A narrative review of the results reported in the included studies on learning outcomes was conducted. The data in the final included studies did not allow for a formal meta-analysis as the studies were not sufficiently homogenous, given the stated research question and the use of different technologies and educational topics.

Results

Study Selection

We identified 15,627 articles from the primary database search. After duplicates were removed, there were 86.63% (13,538/15,627) of articles left for abstract review. Abstracts were screened and, of those 13,538 articles, 179 (1.32%) remained for a full paper review. Of those 179 articles, 150 (83.8%) were excluded, leaving 29 (16.2%) full papers for study inclusion. Details of the study selection process are displayed in Figure 1. In total, 29 RCT studies (N=2722) were included in this review. All studies were conducted in the past 10 years, with most studies (18/29, 62%) published within the past 3 years.

Study Designs

In total, 2722 students participated in the 29 RCTs. Of the 29 articles, 19 (66%) included medical students (1706/2722, 62.67%), 8 (28%) included nursing students (727/2722, 26.71%), and none of the studies involved midwifery students. Approximately 7% (2/29) of studies included both medical and nursing students (289/2722, 10.62%). The following disciplines were used to test the immersive technologies: anatomy, basic clinical and history-taking skills, neurology, respiratory medicine, acute medicine, dermatology, communication skills, internal medicine, and emergency medicine. A full list of the RCTs, basic demographic details, and immersive technology applications included in this review is outlined in Table 1.

Table 1. Randomized controlled trials included in this review of immersive educational tools.

Author	Setting	Application detail	Sample size, N	Purpose	Outcome
Seifert et al [32]	Germany	VP[a] cases (Moodle learning management system)	40	VP—basic clinical skills	Similar levels of long-term knowledge gained; participants assessed the learning experience and the comprehensibility of the seminars as either *very good* or *good*
Wang et al [33]	New Zealand	3D visualizer software (preloaded 3D hologram) on Microsoft HoloLens device	52	Anatomy teaching	There was no difference in knowledge acquisition between groups; only MR[b] group demonstrated higher retention in nominal and spatial types of information; increased engagement in 3DM[c] and MR group
Rossler et al [34]	United States	Virtual Electrosurgery Skill Trainer developed by the National Institutes of Health	20	Fire safety knowledge	No differences in knowledge; intervention group participants were noted to meet performance criteria for their assigned role in their perioperative team
Lombardi et al [35]	United States	Virtual heart activities using physiology software programs (Practice Anatomy Lab, Pearson Education, and Interactive Physiology)	29	Anatomy teaching	Plastic model group achieved significantly higher overall scores on initial and follow-up exams; attitude surveys demonstrated a higher preference for organ dissection
Padilha et al [36]	Portugal	Body Interact (simulation with VPs)	42	Respiratory medicine	Improved knowledge and higher levels of learning satisfaction in the intervention group; no statistically significant differences in self-efficacy perceptions
Blanie et al [37]	France	VP cases—LabForSIMS (simulation center) and a software designer (Interaction Healthcare)	146	Basic clinical skills	No significant educational difference was found; satisfaction and motivation were found to be greater with the use of SG[d]
Liaw et al [38]	Singapore	VP simulation—eRAPIDS, developed at the National University of Singapore	57	Clinical deterioration	No difference in knowledge acquisition; VP was rated positively
Menzel et al [39]	United States	Second Life (Linden Lab) virtual simulation environment (WALD[e] Island)	51	Cultural attitudes	No statistically significant differences between the learning formats
Gananasegaram et al [40]	Canada	Campbell's 3DM of the inner ear—publicly available data sets displayed on Microsoft HoloLens	29	Anatomy teaching	No difference in knowledge acquisition; HG[f] group rated higher for overall effectiveness, ability to convey spatial relationships, and learner engagement and motivation
Liaw et al [41]	Singapore	VR[g] (no details)	198	VP to teach MDT[h] rounds	Increased levels of self-efficacy and attitudes toward interprofessional team care
Moro et al [42]	Australia	Microsoft HoloLens, 3D Studio Max (Autodesk Inc), Unity 3D (Unity Technologies), Vuforia v5 plug-in for Unity (PTC Inc), Samsung Galaxy Tab 3 (Samsung Electronics), Visual Studio v2019	40	Physiology and brain anatomy	No difference in knowledge test scores; significant increase in dizziness using the HoloLens
Stepan et al [26]	United States	VR model of brain anatomy—brain CT[i] scans and MRIs[j], Surgical Theater, Oculus Rift VR system (Oculus VR)	66	Cerebral anatomy	No difference in anatomy knowledge; VR group found learning experience to be significantly more engaging, enjoyable, useful, and motivating
Hu et al [43]	Taiwan	Anatomy Master module of Medical Holodeck	101	Anatomy teaching	Significant improvement in ultrasound task performance and ultrasonographic image identification MCQ[k] tests in the VR group

Author	Setting	Application detail	Sample size, N	Purpose	Outcome
Engum et al [44]	United States	CathSim Intravenous Training System (CathSim) developed by HT Medical (Immersion)	93	Intravenous catheter training	Significant improvement in cognitive gains, student satisfaction, and documentation of the procedure with the traditional laboratory group compared with the computer catheter simulator group
Berg et al [45]	Norway	VR application developed by the authors with hired help for programming (Unity 2018.30f2) and video of the VR features	289	ABCDE basic resuscitation skills	Noninferiority of learning modality; more students in VR group reported liking the way they practiced and that it was a good way to learn; VR group scored high on the System Usability Scale
Kiesewetter et al [46]	Germany	VR learning environment CASUS	142	VP to teach clinical skills	Case formats with a VP did not affect knowledge gain or diagnostic accuracy [46]
Schoeb et al [47]	Germany	Instructions for catheterization displayed on Microsoft HoloLens	164	Catheter training	MR group had significantly better learning outcomes [47]
Noll et al [48]	Germany	AR[l] mobile app, iPhone operating system (iOS, Apple Inc)–based app mArble Derma (m-ARBLE-dermatology)	44	Dermatological teaching	No difference in outcomes between groups [48]
Liaw et al [49]	Singapore	3D virtual hospital developed—CREATIVE	120	Interprofessional skill training	No difference between groups in communication performance scores [49]
Ienghong et al [50]	Thailand	3D USS[m] images played on the downloaded phone app and AR	46	Emergency ultrasound skills	Better performance scores in VR flash card group [50]
Sobocan et al [51]	Slovenia	VP—no detail	34	Internal medicine skills	No difference in exam performance between groups
Kockro et al [52]	Switzerland	Virtual 3DM developed from MRI and CT scans and DextroBeam system (Bracco Advanced Medical Technologies)	169	Neuroanatomy	There were no significant differences in knowledge scores; participants rated the 3D method as superior to 2D teaching methods in four domains: spatial understanding, application in future anatomy classes, effectiveness, and enjoyableness [52]
Nickel et al [53]	Germany	Computer-based TM[n] developed using the open-source Medical Imaging Interaction Toolkit software developed by the German Cancer Research Center	410	Liver anatomy	3D group had higher scores, and participants had a preference for 3D training [53]
Berg et al [54]	Norway	ABCDE resuscitation application developed with help from hired programmers for Unity using Oculus Quest software (Oculus)	289	VP to teach clinical skills	Group self-practice of the ABCDE approach in multiplayer, immersive, interactive VR application was noninferior to practice with physical equipment [54]
Bogomolva et al [55]	The Netherlands	DynamicAnatomy AR application developed at the Department of Anatomy and Embryology at the Leiden University Medical Centre for Innovation	60	Anatomy teaching	No significant differences in knowledge scores [55]
O'Rourke et al [56]	United States	VP model simulation developed with a real patient in real time	60	Clinical skill—breaking bad news	No significant between-group differences on the POMS[o] 2 or salivary cortisol concentration following the SP[p] interaction; students had similar emotional and behavioral responses when delivering bad news to SP or vSP[q] [56]

Author	Setting	Application detail	Sample size, N	Purpose	Outcome
Du et al [57]	Taiwan	3DMs—Autodesk 3DS Max software (Autodesk Media and Entertainment) and Unity 3D developed into a VR gaming system for HTC Vive	18	Anatomy teaching	No significant differences in MCQ scores between groups; VR groups scored highly on the interest, competence, and importance subscales of the IMI[r]; both VR groups considered the system to be fun and beneficial to their learning; MP[s] group reported higher stress levels
Maresky et al [12]	Canada	Using TeraRecon (TeraRecon, Inc), Slicer, and The Body VR: Anatomy Viewer private beta version (The Body VR LLC) together with software provided by Sharecare VR (Sharecare Reality Lab)	42	Cardiac anatomy	VR group scored higher on knowledge
Issleib et al [58]	Germany	VR software developed in cooperation between the University of Hamburg and VIREED[t]—VR-BLS[u] course (using the Laerdal (mannequin)	160	Resuscitation training	Classic BLS training with a seminar and training sessions seemed superior to VR in teaching technical skills [58]

[a]VP: virtual patient.

[b]MR: mixed reality.

[c]3DM: 3D model.

[d]SG: simulation by gaming.

[e]WALD Island named for Lillian Wald, a public health nursing pioneer.

[f]HG: holographic.

[g]VR: virtual reality.

[h]MDT: multidisciplinary team.

[i]CT: computed tomography.

[j]MRI: magnetic resonance imaging.

[k]MCQ: multiple-choice question.

[l]AR: augmented reality.

[m]USS: ultrasound scan.

[n]TM: teaching module.

[o]POMS: Profile of Mood States.

[p]SP: simulated patient.

[q]vSP: virtual simulated patient.

[r]IMI: Intrinsic Motivation Inventory.

[s]MP: multiple-player.

[t]VIREED: Virtual Reality Education (medical virtual reality education platform).

[u]BLS: basic life support.

Study Characteristics

Approximately 76% (22/29) of authors used VR applications, which included virtual simulation scenarios and VPs. Approximately 17% (5/29) of articles used AR applications, which involved using the Microsoft HoloLens headset and mobile phone apps. There were 7% (2/29) of studies that used MR.

Of the 29 articles retrieved, 28 (97%) were from high-income countries and 1 (3%) was from an upper–middle-income country. Most studies originated in the United States (6/29, 21%) and Germany (6/29, 21%), followed by Singapore (3/29, 10%), Canada (2/29, 7%), Norway (2/29, 7%), Taiwan (2/29, 7%), Australia (1/29, 3%), France (1/29, 3%), the Netherlands (1/29, 3%), New Zealand (1/29, 3%), Portugal (1/29, 3%), Slovenia (1/29, 3%), Switzerland (1/29, 3%), Thailand (1/29, 3%), and Turkey (1/29, 3%).

Primary and Secondary Learning Outcomes

Learning outcomes were reported in all studies, including outcomes for both knowledge and the participants' learning experience. Knowledge was assessed via multiple-choice question tests, single best answer tests, general clinical knowledge tests, open-ended style tests, or objective structured clinical examinations. Of the 29 studies, 12 (41%) studies evaluated knowledge using multiple-choice question tests. Approximately 31% (9/29) of studies evaluated knowledge via

general clinical knowledge–based tests that used a variety of validated questionnaires such as the Maastricht Assessment of Simulated Patients questionnaire and the Attitudes toward Poverty scale questionnaire [39,56]. Approximately 10% (3/29) of studies assessed knowledge using open-ended style exam questions. Approximately 3% (1/29) of studies evaluated knowledge via an objective structured clinical examination–based exam. Of the 29 studies, 1 (3%) evaluated knowledge via a single best answer test, 1 (3%) did so via interprofessional skill assessment, and 1 (3%) did not specify the type of evaluation test. Approximately 90% (26/29) of studies reported on satisfaction, attitudes, perceptions, opinions, and general facts regarding the technology used. The learning experience was evaluated in various ways, including validated and nonvalidated instruments on satisfaction, engagement, and perceived learning outcomes. Approximately 3% (1/29) of studies used psychometric testing to evaluate the learning experience [33]. Another study measured salivary cortisol levels before and after the intervention to evaluate whether delivering bad news via a real simulated patient or a virtual simulated patient evoked the same psychological stress [56]. Approximately 66% (19/29) of studies used validated scales to assess the learning experience, namely, the Intrinsic Motivation Inventory [57], General Self-Efficacy Scale, Learner Satisfaction with Simulation Scale [36], and Diagnostic Thinking Inventory [51]. A full list of the validated scales used in the RCTs in this review is provided in Multimedia Appendix 2 [26,33,34,36-38,41,42,45-49,51,54-58]. Approximately 7% (2/29) of studies reported on behaviors as an outcome, and none of the studies reported patient or health care outcomes [51,56]. Only 3% (1/29) of studies reported on adverse health outcomes as part of their secondary outcomes [42].

Study Quality Assessment

The MERSQI scale scores ranged from 10 to 15, with a mean score of 12.64 (SD 1.6). The median was 12.5, and the mode was 13.5. The mean (SD) domain and item scores are illustrated in Multimedia Appendix 3. The mean domain scores were highest for study design (mean 1, SD 0), data analysis (mean 0.7, SD 0.46), and outcomes (mean 0.6, SD 0.21). The lowest-scoring domains included type of data (mean 0.5, SD 0), sampling (mean 0.3, SD 0.15), and validity of the evaluation instrument (mean 0.3, SD 0.27). All articles used an RCT (29/29, 100%) study design, which meant that all studies obtained the highest score in this domain.

Of the 29 studies, 1 (3%) was a double-center RCT, and the remaining 28 (97%) were conducted at a single site. In relation to participant response rate, 93% (27/29) of the studies had a high response rate of >75%.

The authors used appropriate statistical analysis in all studies according to the MERSQI [59]. In relation to the validity of the evaluation instrument, 66% (19/29) of studies used validated instruments and scales to assess learning outcomes, including satisfaction, self-efficacy, engagement, and perceptions of the learning experience.

Finally, according to classification using the criteria by Kirkpatrick [31] in the MERSQI scale, 97% (28/29) of studies assessed knowledge as the primary outcome, and 62% (18/29) of studies used a pre- and postlearning experience knowledge assessment.

Discussion

Principal Findings

The aim of this systematic review was to assess the current educational role of immersive technology in medical, midwifery, and nursing education at the university level compared with traditional learning modalities. Second, the review evaluated the associated learning outcomes of VR, AR, and MR and how they were evaluated. A total of 29 RCTs were selected for this review. The main findings of this study indicate that knowledge gain is equal when using VR, AR, or MR in health care student education compared with traditional learning modalities. In addition, VR, AR, and MR provide an enhanced learning experience based on the secondary outcomes of the studies included in this review [12,26,33,36,38,40,45]. This supports the current evidence that immersive educational technology is a useful and valuable learning tool in medical and nursing preclinical university education. The most common form of immersive technology used was VR. The favored use of VR may be due in part to the widespread availability of cost-effective 3D software and web-based material for the development of anatomy tools [14,60]. Comparators were present in every study and ranged from 2D didactic presentations to textbooks. All the studies retrieved involved medical and nursing students in preclinical university education. Interestingly, there were no studies involving midwifery students. Although there are studies involving midwifery students and immersive technologies in the literature, this review found no RCTs within this domain.

The MERSQI tool provided a standardized approach to evaluating the methodological quality of the studies included in this review [59], which resulted in a moderate level of evidence for these studies. The MERSQI tool also allowed us to classify the learning outcomes of the RCTs clearly, with knowledge identified as the most common primary outcome. Secondary outcomes were identified and included attitudes and opinions related to the learner experience (ie, satisfaction, self-efficacy, engagement, and perceptions of the learning experience). All studies were conducted in upper–middle-income or high-income countries, which may be because of the availability of more funding to support the purchase of such equipment. The articles identified for this review were published within the past 10 years, with most (19/29, 66%) being published in the past 3 years. This highlights the recent rise in interest in immersive technologies for use in preclinical university medical and nursing education. The International Data Corporation has forecasted the shipment of 36.7 million VR headset units by 2023 [61].

Anatomy was the most common topic taught with immersive technologies, which may be because of the ability of immersive technologies to enhance the understanding of complex body processes that cannot be visualized [24]. Subjects such as anatomy and embryology require students to translate 2D images into 3D concepts, which can be a cognitive challenge for those who have difficulty visualizing this transformation of images

[62,63]. Students may also mentally rotate complex structures inaccurately, leaving the mind to fill the gaps in any missing structures [64]. The introduction of immersive technology as an educational tool could automatically standardize this process for students and enhance the understanding of certain subjects such as anatomy and embryology. Similarly, midwifery education requires an understanding of complex concepts, such as fetal development in the womb. Therefore, this would be an important area for future research comparing traditional learning tools with immersive technologies.

VPs were developed in 41% (12/29) of the studies in this review. VPs play a key role in learning basic clinical skills for both nursing and medical education. In addition, they create an environment for repeated practice in a safe space without harm to patients [65,66]. Traditionally, clinical education relies on practicing diagnostic, therapeutic, and procedural skills on live patients. This can be difficult to balance on a day-to-day basis because of the fast-paced nature of medicine, time constraints for clinical teaching in busy wards, and competition between students [67]. In addition, over the past year, the experience of the global COVID-19 pandemic has vastly reduced access to bedside teaching for nursing, midwifery, and medical students [68]. Therefore, the availability and development of VPs and immersive learning environments can play a key role in the future as an adjunct to developing clinical skills for students at all levels and at any time [61].

The primary outcome in most studies (28/29, 97%) was knowledge. This review demonstrated that immersive technology is equally effective in knowledge gained by the student and, in some studies, reflects a higher level of knowledge retention. Regarding secondary outcomes concerning the learning experience, all studies reported overall positivity and higher satisfaction in learning, self-efficacy, and engagement with immersive technologies. Moreover, this review revealed the heterogeneity of tools and instruments used to evaluate secondary learning outcomes such as student satisfaction, self-efficacy, engagement, and learning perceptions. Several studies developed their own Likert-style assessment scales, whereas others used and adapted previously developed and validated scales for the assessment of secondary learning outcomes. According to the Association for Medical Education in Europe, “assessment tools selected should be valid, reliable, practical and have an appropriate impact on student learning” [69]. Perhaps this highlights the need for a standardized, validated instrument to be designed specifically for immersive technology in education for the evaluation of learning outcomes related to the students’ learning experience. For example, in simulation, the National League of Nursing developed a validated standardized scale to evaluate the use of simulation and student learning experience with simulation activity, including students’ satisfaction and self-confidence, perceived learning, and engagement [36]. Therefore, the development of a standardized instrument to evaluate learning outcomes, such as satisfaction, self-confidence, self-efficacy, and engagement, for immersive educational tools may be beneficial so that future research may be better informed by a more uniform approach to assessing learning outcomes.

Strengths and Limitations

This systematic review provides an up-to-date review on the learning outcomes of immersive technologies in university-level medical and nursing education. This study illustrates important findings on the use of immersive technologies that will provide a foundation for future research in this area. Knowledge gain with immersive technologies was shown in this review to be equal to or greater than that of traditional modalities, thus providing an evidence base for future curriculum developers and researchers alike to implement these immersive tools into university programs. A comprehensive search strategy and robust methodology support the strengths of this study. The use of the validated MERSQI tool to assess the study design of the included studies also adds to the strength of our study design.

Most of the reviewed studies developed their own bespoke immersive educational tool specific to their needs, with VR being primarily used. The favored use of VR may be due in part to the widespread availability of cost-effective 3D software and web-based material for the development of anatomy tools [14,60]. However, the development of these tools has financial and resource implications, including the time spent developing content for these devices. With the increased amount of content and material, there may be an opportunity to develop a universal platform that makes content sharable and available worldwide for health care education. This may reduce the barriers to accepting and implementing this technology in health care education.

Nevertheless, we recognize that there were also some limitations to this study. Common biases exist within the methodology section, including the study eligibility criteria, identification and selection of studies, data extraction, and study appraisal. Predefined search criteria and inclusion criteria were set out in the published protocol, which aimed to reduce these biases. Within the published literature, there is heterogeneity in how VR is defined, requiring us to confine VR to include the use of a headset, providing only an immersive experience, as opposed to a 3D visualization on a computer screen. Therefore, this may have resulted in the exclusion of studies involving a 3D animation or model on a computer screen that did not place the learner in an immersive environment. Furthermore, because of the initial large database of articles retrieved, we decided to only include RCTs, as they are regarded as the highest-quality study type and also included a comparator that was a traditional learning modality. This process was conducted by 2 independent reviewers (GR and SC), and full-text inclusion was dependent on agreement by both reviewers. During the study retrieval process, bias may have occurred because of the unavailability of some studies. Authors were contacted in cases of unavailable data; however, this may have led to data availability bias and unrepresented data. This study included only preclinical university students; therefore, the value of immersive technology in the postqualification setting is unknown.

Conclusions

In conclusion, this systematic review demonstrates that VR, AR, and MR are beneficial educational tools in preclinical medical and nursing university education. Immersive technology is equally effective in teaching and increases learner satisfaction,

self-confidence, and engagement. However, further research is required to explore the role of VR, AR, and MR in midwifery education. With the increasing availability of cost-effective immersive tools, the use of immersive technologies in health care student education is potentially very valuable.

Authors' Contributions

This study was part of a thesis by GR for an MD in Medical Education, University College Dublin, Ireland. GR and SC were involved in the writing of the study protocol drafts, development of the study design, data analysis, and writing of the manuscript. AR was involved in the data analysis, writing, and review of the manuscript. MH was a study supervisor and was involved in the review of the manuscript. EM was also a study supervisor and reviewed the manuscript. FMA was a study supervisor and was involved in the study design, writing, and review of the study protocol and manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Comprehensive list of search terms developed for the search strategy.
[PDF File (Adobe PDF File), 22 KB - jmir_v24i2e30082_app1.pdf]

Multimedia Appendix 2
Validated scales used for evaluation.
[PDF File (Adobe PDF File), 19 KB - jmir_v24i2e30082_app2.pdf]

Multimedia Appendix 3
Medical Education Research Study Quality Instrument domain and item scores for the 29 studies.
[PDF File (Adobe PDF File), 32 KB - jmir_v24i2e30082_app3.pdf]

References

1. Peterson DC, Mlynarczyk GS. Analysis of traditional versus three-dimensional augmented curriculum on anatomical learning outcome measures. Anat Sci Educ 2016;9(6):529-536 [FREE Full text] [doi: 10.1002/ase.1612] [Medline: 27078503]
2. Jin J, Bridges SM. Educational technologies in problem-based learning in health sciences education: a systematic review. J Med Internet Res 2014;16(12):e251 [FREE Full text] [doi: 10.2196/jmir.3240] [Medline: 25498126]
3. Mangina E. 3D learning objects for augmented/virtual reality educational ecosystems. In: 2017 23rd International Conference on Virtual System & Multimedia.: IEEE; 2017 Presented at: VSMM'17; Octover 31-November 4, 2017; Dublin p. 1-6. [doi: 10.1109/vsmm.2017.8346266]
4. Bell JT, Fogler HS. The investigation and application of virtual reality as an educational tool. In: 1995 American Society of Engineering Education Annual Conference. 1995 Presented at: ASEE'95; June 25-28, 1995; Anahein.
5. Salzman MC, Dede C, Loftin RB, Chen J. A model for understanding how virtual reality aids complex conceptual learning. Presence (Camb) 1999;8(3):293-316. [doi: 10.1162/105474699566242]
6. Ellaway R, Masters K. AMEE guide 32: e-learning in medical education part 1: learning, teaching and assessment. Med Teach 2008;30(5):455-473. [doi: 10.1080/01421590802108331] [Medline: 18576185]
7. Kapralos B, Uribe-Quevedo A, Dubrowski A. Immersive technologies for medical education. In: Lee N, editor. Encyclopedia of computer graphics and games. Ney York City: Springer International Publishing; 2018:1-8.
8. Anderson EF, McLoughlin L, Liarokapis F, Peters CE, Petridis P, de Freitas S. Developing serious games for cultural heritage: a state-of-the-art review. Virtual Real 2010;14(4):255-275. [doi: 10.1007/s10055-010-0177-3]
9. Hussein M, Nätterdal C. The benefits of virtual reality in education- a comparision study. Göteborgs Universitet. Gothenburg: University of Gothenburg; 2015. URL: http://hdl.handle.net/2077/39977 [accessed 2020-12-15]
10. Pantelidis P, Chorti A, Papagiouvanni I, Paparoidamis G, Drosos C, Panagiotakopoulos T, et al. Virtual and augmented reality in medical education. In: Tsoulfas G, editor. Medical and Surgical Education-Past, Present and Future. London: IntechOpen; 2017:77-97.
11. Nicola S, Stoicu-Tivadar L. Mixed reality supporting modern medical education. Stud Health Technol Inform 2018;255:242-246. [Medline: 30306945]
12. Maresky HS, Oikonomou A, Ali I, Ditkofsky N, Pakkal M, Ballyk B. Virtual reality and cardiac anatomy: exploring immersive three‐dimensional cardiac imaging, a pilot study in undergraduate medical anatomy education. Clin Anat 2019;32(2):238-243. [doi: 10.1002/ca.23292] [Medline: 30295333]
13. Moro C, Štromberga Z, Raikos A, Stirling A. The effectiveness of virtual and augmented reality in health sciences and medical anatomy. Anat Sci Educ 2017;10(6):549-559. [doi: 10.1002/ase.1696] [Medline: 28419750]

14. Nicholson DT, Chalk C, Funnell WR, Daniel SJ. Can virtual reality improve anatomy education? A randomised controlled study of a computer-generated three-dimensional anatomical ear model. Med Educ 2006;40(11):1081-1087. [doi: 10.1111/j.1365-2929.2006.02611.x] [Medline: 17054617]
15. Hansen MM. Versatile, immersive, creative and dynamic virtual 3-D healthcare learning environments: a review of the literature. J Med Internet Res 2008;10(3):e26. [doi: 10.2196/jmir.1051] [Medline: 18762473]
16. Liyanagunawardena TR, Williams SA. Massive open online courses on health and medicine: review. J Med Internet Res 2014;16(8):e191 [FREE Full text] [doi: 10.2196/jmir.3439] [Medline: 25123952]
17. Paim CP, Goldmeier S. Development of an educational game to set up surgical instruments on the mayo stand or back table: applied research in production technology. JMIR Serious Games 2017;5(1):e1 [FREE Full text] [doi: 10.2196/games.6048] [Medline: 28073736]
18. Graafland M, Schraagen JM, Schijven MP. Systematic review of serious games for medical education and surgical skills training. Br J Surg 2012;99(10):1322-1330 [FREE Full text] [doi: 10.1002/bjs.8819] [Medline: 22961509]
19. Aksoy E. Comparing the effects on learning outcomes of tablet-based and virtual reality-based serious gaming modules for basic life support training: randomized trial. JMIR Serious Games 2019;7(2):e13442 [FREE Full text] [doi: 10.2196/13442] [Medline: 31042153]
20. Car J, Carlstedt-Duke J, Tudor Car L, Posadzki P, Whiting P, Zary N, Digital Health Education Collaboration. Digital education in health professions: the need for overarching evidence synthesis. J Med Internet Res 2019;21(2):e12913 [FREE Full text] [doi: 10.2196/12913] [Medline: 30762583]
21. Akl EA, Pretorius RW, Sackett K, Erdley WS, Bhoopathi PS, Alfarah Z, et al. The effect of educational games on medical students' learning outcomes: a systematic review: BEME guide No 14. Med Teach 2010;32(1):16-27. [doi: 10.3109/01421590903473969] [Medline: 20095770]
22. De Gagne JC, Oh J, Kang J, Vorderstrasse AA, Johnson CM. Virtual worlds in nursing education: a synthesis of the literature. J Nurs Educ 2013;52(7):391-396. [doi: 10.3928/01484834-20130610-03] [Medline: 23755942]
23. Tang KS, Cheng DL, Mi E, Greenberg PB. Augmented reality in medical education: a systematic review. Can Med Educ J 2020;16(11):e81-e96. [doi: 10.36834/cmej.61705] [Medline: 32215146]
24. Zhao J, Xu X, Jiang H, Ding Y. The effectiveness of virtual reality-based technology on anatomy teaching: a meta-analysis of randomized controlled studies. BMC Med Educ 2020;20(1):127. [doi: 10.1186/s12909-020-1994-z] [Medline: 32334594]
25. Forgione A, Guraya S. The cutting-edge training modalities and educational platforms for accredited surgical training: a systematic review. J Res Med Sci 2017;22:51. [doi: 10.4103/jrms.jrms_809_16] [Medline: 28567070]
26. Stepan K, Zeiger J, Hanchuk S, Del Signore A, Shrivastava R, Govindaraj S, et al. Immersive virtual reality as a teaching tool for neuroanatomy. Int Forum Allergy Rhinol 2017;7(10):1006-1013. [doi: 10.1002/alr.21986] [Medline: 28719062]
27. Alfalah SF, Falah JF, Muhaidat NM, Hudaib A, Koshebye D, AlHourani S. Virtual reality learning environment in embryology education. Int J Educ Pedagogical Sci 2019;13(12):1461-1464. [doi: 10.5281/zenodo.3593206]
28. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009;339:b2535. [doi: 10.1136/bmj.b2535] [Medline: 19622551]
29. Higgins JT, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021). Cochrane. 2021. URL: https://training.cochrane.org/handbook [accessed 2020-10-02]
30. Cook DA, Reed DA. Appraising the quality of medical education research methods: the medical education research study quality instrument and the Newcastle-Ottawa scale-education. Acad Med 2015;90(8):1067-1076. [doi: 10.1097/ACM.0000000000000786] [Medline: 26107881]
31. Kirkpatrick DL, Kirkpatrick JD. Implementing the four levels: a practical guide for effective evaluation of training programs. Oakland: Berrett-Koehler Publishers; 2009.
32. Seifert L, Manap A, Sterz J, Gerlach F, Sader R. A comparison between virtual patient and peer-assisted learning in teaching basic medical knowledge and skills. Electron J e-Learn 2020;18(1):40-56. [doi: 10.34190/EJEL.20.18.1.004]
33. Wang C, Daniel BK, Asil M, Khwaounjoo P, Cakmak YO. A randomised control trial and comparative analysis of multi-dimensional learning tools in anatomy. Sci Rep 2020;10(1):6120. [doi: 10.1038/s41598-020-62855-6] [Medline: 32273563]
34. Rossler KL, Sankaranarayanan G, Duvall A. Acquisition of fire safety knowledge and skills with virtual reality simulation. Nurse Educ 2019;44(2):88-92 [FREE Full text] [doi: 10.1097/NNE.0000000000000551] [Medline: 29847356]
35. Lombardi SA, Hicks RE, Thompson KV, Marbach-Ad G. Are all hands-on activities equally effective? Effect of using plastic models, organ dissections, and virtual dissections on student learning and perceptions. Adv Physiol Educ 2014;38(1):80-86 [FREE Full text] [doi: 10.1152/advan.00154.2012] [Medline: 24585474]
36. Padilha JM, Machado PP, Ribeiro A, Ramos J, Costa P. Clinical virtual simulation in nursing education: randomized controlled trial. J Med Internet Res 2019;21(3):e11529. [doi: 10.2196/11529] [Medline: 30882355]
37. Blanié A, Amorim MA, Benhamou D. Comparative value of a simulation by gaming and a traditional teaching method to improve clinical reasoning skills necessary to detect patient deterioration: a randomized study in nursing students. BMC Med Educ 2020;20(1):53 [FREE Full text] [doi: 10.1186/s12909-020-1939-6] [Medline: 32075641]

38. Liaw SY, Chan SW, Chen FG, Hooi SC, Siau C. Comparison of virtual patient simulation with mannequin-based simulation for improving clinical performances in assessing and managing clinical deterioration: randomized controlled trial. J Med Internet Res 2014;16(9):e214 [FREE Full text] [doi: 10.2196/jmir.3322] [Medline: 25230684]
39. Menzel N, Willson LH, Doolen J. Effectiveness of a poverty simulation in Second Life®: changing nursing student attitudes toward poor people. Int J Nurs Educ Scholarsh 2014;11(1):39-45. [doi: 10.1515/ijnes-2013-0076] [Medline: 24615491]
40. Gnanasegaram JJ, Leung R, Beyea JA. Evaluating the effectiveness of learning ear anatomy using holographic models. J Otolaryngol Head Neck Surg 2020;49(1):63. [doi: 10.1186/s40463-020-00458-x] [Medline: 32814593]
41. Liaw SY, Tan KK, Wu LT, Tan SC, Choo H, Yap J, et al. Finding the right blend of technologically enhanced learning environments: randomized controlled study of the effect of instructional sequences on interprofessional learning. J Med Internet Res 2019;21(5):e12537 [FREE Full text] [doi: 10.2196/12537] [Medline: 31140432]
42. Moro C, Phelps C, Redmond P, Stromberga Z. HoloLens and mobile augmented reality in medical and health science education: a randomised controlled trial. Br J Educ Technol 2021;52(2):680-694. [doi: 10.1111/bjet.13049]
43. Hu KC, Salcedo D, Kang YN, Lin CW, Hsu CW, Cheng CY, et al. Impact of virtual reality anatomy training on ultrasound competency development: a randomized controlled trial. PLoS One 2020;15(11):e0242731 [FREE Full text] [doi: 10.1371/journal.pone.0242731] [Medline: 33227037]
44. Engum SA, Jeffries P, Fisher L. Intravenous catheter training system: computer-based education versus traditional learning methods. Am J Surg 2003;186(1):67-74. [doi: 10.1016/s0002-9610(03)00109-0] [Medline: 12842753]
45. Berg H, Steinsbekk A. Is individual practice in an immersive and interactive virtual reality application non-inferior to practicing with traditional equipment in learning systematic clinical observation? A randomized controlled trial. BMC Med Educ 2020;20(1):123. [doi: 10.1186/s12909-020-02030-7] [Medline: 32326948]
46. Kiesewetter J, Sailer M, Jung VM, Schönberger R, Bauer E, Zottmann JM, et al. Learning clinical reasoning: how virtual patient case format and prior knowledge interact. BMC Med Educ 2020;20(1):73 [FREE Full text] [doi: 10.1186/s12909-020-1987-y] [Medline: 32171297]
47. Schoeb DS, Schwarz J, Hein S, Schlager D, Pohlmann PF, Frankenschmidt A, et al. Mixed reality for teaching catheter placement to medical students: a randomized single-blinded, prospective trial. BMC Med Educ 2020;20(1):510 [FREE Full text] [doi: 10.1186/s12909-020-02450-5] [Medline: 33327963]
48. Noll C, von Jan U, Raap U, Albrecht UV. Mobile augmented reality as a feature for self-oriented, blended learning in medicine: randomized controlled trial. JMIR Mhealth Uhealth 2017;5(9):e139 [FREE Full text] [doi: 10.2196/mhealth.7943] [Medline: 28912113]
49. Liaw SY, Ooi SW, Rusli KD, Lau TC, Tam WW, Chua WL. Nurse-physician communication team training in virtual reality versus live simulations: randomized controlled trial on team communication and teamwork attitudes. J Med Internet Res 2020;22(4):e17279 [FREE Full text] [doi: 10.2196/17279] [Medline: 32267235]
50. Ienghong K, Kotruchin P, Tangpaisarn T, Apiratwarakul K. Practical emergency ultrasound flashcards with augmented reality in teaching point-of-care ultrasound in ER. Open Access Maced J Med Sci 2021;9(E):39-42. [doi: 10.3889/oamjms.2021.5566]
51. Sobocan M, Turk N, Dinevski D, Hojs R, Pecovnik Balon B. Problem-based learning in internal medicine: virtual patients or paper-based problems? Intern Med J 2017;47(1):99-103. [doi: 10.1111/imj.13304] [Medline: 27800653]
52. Kockro RA, Amaxopoulou C, Killeen T, Wagner W, Reisch R, Schwandt E, et al. Stereoscopic neuroanatomy lectures using a three-dimensional virtual reality environment. Ann Anat 2015;201:91-98. [doi: 10.1016/j.aanat.2015.05.006] [Medline: 26245861]
53. Nickel F, Hendrie JD, Bruckner T, Kowalewski KF, Kenngott HG, Müller-Stich BP, et al. Successful learning of surgical liver anatomy in a computer-based teaching module. Int J Comput Assist Radiol Surg 2016;11(12):2295-2301. [doi: 10.1007/s11548-016-1354-y] [Medline: 26872808]
54. Berg H, Steinsbekk A. The effect of self-practicing systematic clinical observations in a multiplayer, immersive, interactive virtual reality application versus physical equipment: a randomized controlled trial. Adv Health Sci Educ Theory Pract 2021;26(2):667-682 [FREE Full text] [doi: 10.1007/s10459-020-10019-6] [Medline: 33511505]
55. Bogomolova K, van der Ham IJ, Dankbaar ME, van den Broek WW, Hovius SE, van der Hage JA, et al. The effect of stereoscopic augmented reality visualization on learning anatomy and the modifying effect of visual-spatial abilities: a double-center randomized controlled trial. Anat Sci Educ 2020;13(5):558-567. [doi: 10.1002/ase.1941] [Medline: 31887792]
56. O'Rourke SR, Branford KR, Brooks TL, Ives LT, Nagendran A, Compton SN. The emotional and behavioral impact of delivering bad news to virtual versus real standardized patients: a pilot study. Teach Learn Med 2019;32(2):139-149. [doi: 10.1080/10401334.2019.1652180] [Medline: 31437006]
57. Du YC, Fan SC, Yang LC. The impact of multi-person virtual reality competitive learning on anatomy education: a randomized controlled study. BMC Med Educ 2020;20(1):343. [doi: 10.1186/s12909-020-02155-9] [Medline: 33081774]
58. Issleib M, Kromer A, Pinnschmidt HO, Süss-Havemann C, Kubitz JC. Virtual reality as a teaching method for resuscitation training in undergraduate first year medical students: a randomized controlled trial. Scand J Trauma Resusc Emerg Med 2021;29(1):27 [FREE Full text] [doi: 10.1186/s13049-021-00836-y] [Medline: 33526042]
59. Reed DA, Cook DA, Beckman TJ, Levine RB, Kern DE, Wright SM. Association between funding and quality of published medical education research. JAMA 2007;298(9):1002-1009. [doi: 10.1001/jama.298.9.1002] [Medline: 17785645]

60. Hampton BS, Sung VW. Improving medical student knowledge of female pelvic floor dysfunction and anatomy: a randomized trial. Am J Obstet Gynecol 2010;202(6):601.e1-601.e8. [doi: 10.1016/j.ajog.2009.08.038] [Medline: 20430356]
61. Pears M, Yiasemidou M, Ismail MA, Veneziano D, Biyani CS. Role of immersive technologies in healthcare education during the COVID-19 epidemic. Scott Med J 2020;65(4):112-119. [doi: 10.1177/0036933020956317] [Medline: 32878575]
62. Marsh KR, Giffin BF, Lowrie Jr DJ. Medical student retention of embryonic development: impact of the dimensions added by multimedia tutorials. Anat Sci Ed 2008;1(6):252-257. [doi: 10.1002/ase.56] [Medline: 19109854]
63. Brewer DN, Wilson TD, Eagleson R, de Ribaupierre S. Evaluation of neuroanatomical training using a 3D visual reality model. Stud Health Technol Inform 2012;173:85-91. [Medline: 22356963]
64. Liesefeld HR, Fu X, Zimmer HD. Fast and careless or careful and slow? Apparent holistic processing in mental rotation is explained by speed-accuracy trade-offs. J Exp Psychol Learn Mem Cogn 2015;41(4):1140-1151. [doi: 10.1037/xlm0000081] [Medline: 25528084]
65. Persky S, McBride CM. Immersive virtual environment technology: a promising tool for future social and behavioral genomics research and practice. Health Commun 2009;24(8):677-682. [doi: 10.1080/10410230903263982] [Medline: 20183376]
66. Chen FQ, Leng YF, Ge JF, Wang DW, Li C, Chen B, et al. Effectiveness of virtual reality in nursing education: meta-analysis. J Med Internet Res 2020;22(9):e18290. [doi: 10.2196/18290] [Medline: 32930664]
67. Spencer J. Learning and teaching in the clinical environment. BMJ 2003;326(7389):591-594 [FREE Full text] [doi: 10.1136/bmj.326.7389.591] [Medline: 12637408]
68. Remtulla R. The present and future applications of technology in adapting medical education amidst the COVID-19 pandemic. JMIR Med Educ 2020;6(2):e20190 [FREE Full text] [doi: 10.2196/20190] [Medline: 32634107]
69. Shumway JM, Harden RM, Association for Medical Education in Europe. AMEE guide No. 25: the assessment of learning outcomes for the competent and reflective physician. Med Teach 2003;25(6):569-584. [doi: 10.1080/0142159032000151907] [Medline: 15369904]

Abbreviations

AR: augmented reality
MERSQI: Medical Education Research Study Quality Instrument
MR: mixed reality
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO: International Prospective Register of Systematic Reviews
RCT: randomized controlled trial
VP: virtual patient
VR: virtual reality

Edited by R Kukafka; submitted 30.04.21; peer-reviewed by R Lundin, A Roy, S Mann; comments to author 28.06.21; revised version received 11.07.21; accepted 26.10.21; published 01.02.22.

Please cite as:
Ryan GV, Callaghan S, Rafferty A, Higgins MF, Mangina E, McAuliffe F
Learning Outcomes of Immersive Technologies in Health Care Student Education: Systematic Review of the Literature
J Med Internet Res 2022;24(2):e30082
URL: https://www.jmir.org/2022/2/e30082
doi:10.2196/30082
PMID:35103607

Review

Effectiveness of eHealth and mHealth Interventions Supporting Children and Young People Living With Juvenile Idiopathic Arthritis: Systematic Review and Meta-analysis

Sonia Butler[1*], BN (Hons), GradCertTertiaryEd, MEd; Dean Sculley[1*], BSc (Hons), MSc, PhD; Derek Santos[2*], BSc (Hons), MSc, PhD; Antoni Fellas[3*], BHSc (Hons), PhD; Xavier Gironès[4*], BSc (Hons), PhD; Davinder Singh-Grewal[5,6,7,8*], MBBS, MMedSci, PhD, FRACP; Andrea Coda[3,9*], BSc (Hons), PhD

[1]School of Bioscience and Pharmacy, College of Health, Medicine and Wellbeing, University of Newcastle, Ourimbah, Australia

[2]School of Health Sciences, Queen Margaret University, Edinburgh, United Kingdom

[3]School of Health Sciences, College of Health, Medicine and Wellbeing, University of Newcastle, Callaghan, Australia

[4]University of Vic-Central University of Catalonia, Manresa, Spain

[5]Department of Rheumatology, Sydney Children's Hospitals Network, Randwick and Westmead, Sydney, Australia

[6]Department of Rheumatology, John Hunter Children's Hospital, Newcastle, Australia

[7]School of Women's and Children's Health, University of New South Wales, Sydney, Australia

[8]Discipline of Child and Adolescent Health, University of Sydney, Sydney, Australia

[9]Priority Research Centre Health Behaviour, Hunter Medical Research Institute, Newcastle, Australia

[*]all authors contributed equally

Corresponding Author:
Sonia Butler, BN (Hons), GradCertTertiaryEd, MEd
School of Bioscience and Pharmacy
College of Health, Medicine and Wellbeing
University of Newcastle
10 Chittaway Rd
Ourimbah, 2258
Australia
Phone: 61 421945914
Email: sonia.butler@newcastle.edu.au

Abstract

Background: Juvenile idiopathic arthritis (JIA) management aims to promote remission through timely, individualized, well-coordinated interdisciplinary care using a range of pharmacological, physical, psychological, and educational interventions. However, achieving this goal is workforce-intensive. Harnessing the burgeoning eHealth and mobile health (mHealth) interventions could be a resource-efficient way of supplementing JIA management.

Objective: This systematic review aims to identify the eHealth and mHealth interventions that have been proven to be effective in supporting health outcomes for children and young people (aged 1-18 years) living with JIA.

Methods: We systematically searched 15 databases (2018-2021). Studies were eligible if they considered children and young people (aged 1-18 years) diagnosed with JIA, an eHealth or mHealth intervention, any comparator, and health outcomes related to the used interventions. Independently, 2 reviewers screened the studies for inclusion and appraised the study quality using the Downs and Black (modified) checklist. Study outcomes were summarized using a narrative, descriptive method and, where possible, combined for a meta-analysis using a random-effects model.

Results: Of the 301 studies identified in the search strategy, 15 (5%) fair-to-good–quality studies met the inclusion criteria, which identified 10 interventions for JIA (age 4-18.6 years). Of these 10 interventions, 5 (50%) supported symptom monitoring by capturing real-time data using health applications, electronic diaries, or web-based portals to monitor pain or health-related quality of life (HRQoL). Within individual studies, a preference was demonstrated for real-time pain monitoring over recall pain assessments because of a peak-end effect, improved time efficiency (P=.002), and meeting children's and young people's HRQoL needs (P<.001) during pediatric rheumatology consultations. Furthermore, 20% (2/10) of interventions supported physical activity promotion using a web-based program or a wearable activity tracker. The web-based program exhibited a moderate effect, which

increased endurance time, physical activity levels, and moderate to vigorous physical activity (standardized mean difference [SMD] 0.60, SD 0.02-1.18; I^2=79%; P=.04). The final 30% (3/10) of interventions supported self-management development through web-based programs, or apps, facilitating a small effect, reducing pain intensity (SMD −0.14, 95% CI −0.43 to 0.15; I^2=53%; P=.33), and increasing disease knowledge and self-efficacy (SMD 0.30, 95% CI 0.03-0.56; I^2=74%; P=.03). These results were not statistically significant. No effect was seen regarding pain interference, HRQoL, anxiety, depression, pain coping, disease activity, functional ability, or treatment adherence.

Conclusions: Evidence that supports the inclusion of eHealth and mHealth interventions in JIA management is increasing. However, this evidence needs to be considered cautiously because of the small sample size, wide CIs, and moderate to high statistical heterogeneity. More rigorous research is needed on the longitudinal effects of real-time monitoring, web-based pediatric rheumatologist–children and young people interactions, the comparison among different self-management programs, and the use of wearable technologies as an objective measurement for monitoring physical activity before any recommendations that inform current practice can be given.

(J Med Internet Res 2022;24(2):e30457) doi:10.2196/30457

KEYWORDS

eHealth; mobile health; mHealth; juvenile idiopathic arthritis; pediatric; effectiveness; pain; physical activity; health-related quality of life; self-management; education; mobile phone

Introduction

Background

Juvenile idiopathic arthritis (JIA) is the most common rheumatic disease in pediatric populations [1]. Early diagnosis and active treatment are essential for maintaining physical function and psychological well-being. The treatments aim to control the disease, promote clinical remission, and prevent long-term disability [2-5]. However, to achieve these goals, the management of JIA should be multifactorial [6]. Pediatric-specific issues need tending, such as the use of antirheumatic medications in children and young people, growth retardation, pain and coping, school attendance, psychosocial functioning, dealing with parents, and, in the adolescent years, preparing for the transition to adult care [4,7,8]. For good reason, children and young people need to be closely monitored and supported by specialized rheumatology centers that provide interdisciplinary care using a range of pharmacological, physical, psychological, and educational interventions [6,9-12]. However, several barriers have been identified that hinder this current model of support, delaying the delivery of timely, individualized, and well-coordinated care.

There is an inadequate number of experienced pediatric rheumatologists (PRs) to meet demand and oversee care [7,13-18]. This has resulted in long waiting lists, the centralization of services into tertiary children's hospitals, and the need for many children and young people to travel long distances to access pediatric rheumatology centers [2,18] or care being delivered by a primary health care provider with no pediatric rheumatology training [2,5,14,18-20]. The World Forum on Rheumatic and Musculoskeletal Diseases clearly states that poor access to health care services can significantly impede diagnosis, appropriate treatment, and health outcomes [14], highlighting the need to overcome these barriers in the delivery of JIA management.

In addition, to achieve optimal health outcomes, children and young people need to adhere to their prescribed treatment plan [21,22], and parents need to support treatment recommendations [2,22]. However, suboptimal rates of adherence are commonly reported [22-24]. For example, a literature review of children and young people with chronic rheumatoid disease reported medication adherence rates as low as 38% and physical activity adherence rates of 40%, particularly during adolescence [22]. The primary reasons included the complexities of chronic disease management and medication schedules, time-consuming nonpharmacological treatments, lack of disease knowledge, and low satisfaction with the health care team [22]. These reasons are not surprising, as a recent systematic review identified 70 studies in which health information was inappropriately tailored to children and young people and their parents' level of health literacy, increasing their concerns and uncertainties about their condition, treatment options, and shared care decisions [25].

For JIA specifically, further reasons for nonadherence vary across treatment modalities [24] as follows: for oral medications, forgetfulness, taste, and long-term side-effects; for parenteral medications (injectables and infusions), pain and side-effects; and for physical and occupational therapy, forgetfulness, pain, and therapy not being considered necessary [24]. Fortunately, all these reasons are modifiable.

To uphold the expectations of rheumatology care, children and young people should be empowered to take an active role in their disease management by being provided opportunities to improve their health literacy and develop good self-management skills [2,10,25], particularly when considering the long-term benefits these skills will have across their life span. The development of self-management skills is also important as parents only hold a surrogate role in children's and young people's health care decisions; therefore, children and young people need to be prepared for their transition from pediatric to adult health care services [26].

A resource-efficient way of supplementing JIA management and the development of self-management skills could be to harness the burgeoning eHealth and mobile health (mHealth) interventions [27]. eHealth is described by the World Health Organization as an activity that delivers health-related information, resources, and services through electronic

technology and the internet [28]. mHealth is described as a subbranch of eHealth [28] that uses wireless technology to rapidly uptake, process, and communicate information to support health system efficiency and patient outcomes [29].

The number of studies exploring the potential of eHealth and mHealth interventions for chronic disease management is rapidly increasing. However, most are still at an early stage of development and are limited in their scientific rigor [30-35]; most have been conducted with adults rather than children and young people [31,35], which is interesting, considering that children and young people are experienced users of this form of technology and more likely to use a digital intervention or health app [36-38]. In fact, a recent systematic review identified that children and young people use the internet for at least 1 to 4 hours a day (9438/10,974, 86%) [37] and some type of app every day (719/719, 100%) [39]. Higher rates of use have also been reported for children and young people living with JIA who are at risk of poor psychosocial functioning compared with their peers (>1 hour a day) [36].

However, concerns have been raised about how children and young people use the internet [37,40]. Studies have established that children and young people have poor internet-searching skills, tend to use a 1-word search strategy, briefly skim through search-engine result pages [40], and lack the ability to appraise quality [33,36]. This limits their capacity to find high-quality, personally relevant health information and potentially exposes them to incorrect material [36] or results in them turning to apps and platforms not specifically developed as health resources such as YouTube, Tumblr, and Instagram [41].

Johnson et al [36] believe that for pediatric services to better support the needs of children and young people living with chronic illness, they need to be provided with accessible, developmentally appropriate, and high-quality health-related information. Children and young people with JIA (n=134) agreed, particularly those with low health-related quality of life (HRQoL), expressing an interest in being provided with supportive internet-based interventions [36]. In addition, children and young people participating in feasibility and usability studies and reporting on the delivery of health messages [42], exercise programs [43], symptom monitoring [44,45], and disease management [35,36,44] have also reported high levels of acceptability [42-44], usefulness [35], and satisfaction [43] when using these interventions. However, personal, technical, and device-related barriers have also been identified, which hinder their use [46]. Understandably, before a health care professional can prescribe a digital intervention, it has been suggested that they need at least 3 published papers demonstrating the intervention's effectiveness [47] to see whether the intervention works in a real-world setting [48].

Definition of Children and Young People

Internationally, pediatric services cater to children aged 0 to 12 years [49], and adolescents up to the age of 18 years (mean 18.7, SD 2.6 years), before they are transferred to adult services [50]. In this review, we use the term "children and young people" to broadly include all individuals in the age range of 1 to 18 years. We exclude neonates and infants (<1 year) [51].

Aim and Rationale

This systematic review aims to identify what eHealth and mHealth interventions have proven to be effective in supporting health outcomes for children and young people (aged 1-18 years) living with JIA. We anticipate that this review may aid the clinical use of eHealth and mHealth interventions and their integration in arthritis management.

Methods

Overview

This systematic review complies with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [52]. Before commencement of this review, a protocol for this review was registered on PROSPERO (CRD42018108985) [53]. Protocol questions 1, 2, and 4 were presented in a previous study [46], whereas question 3 is presented in this review [53]:

1. What types of eHealth and mHealth interventions have been used to investigate the health care of children and young people diagnosed with JIA?
2. What is the usability of eHealth and mHealth interventions for children and young people diagnosed with JIA?
3. What eHealth and mHealth interventions have proven to be effective in helping children and young people diagnosed with JIA?
4. Are the existing eHealth and mHealth interventions cost-effective for pediatric rheumatology?

Eligibility Criteria

Search Strategy

The search terms in this review were developed by SB after initially searching the National Center for Biotechnology Information Medical Subject Heading terms (Multimedia Appendix 1) [54]. The search terms were adapted to suit 15 health databases with the aim of gaining a broad range of interdisciplinary literature. These databases included MEDLINE or PubMed, Cochrane Library, Joanna Briggs Institute, AMED, CINAHL complete, Embase, JAMA, Informit Health, ProQuest database, PsycINFO, IEEE Xplore, SAGE Publishing, ScienceDirect, Scopus, and Web of Science. The search was conducted in October 2018 and July 2021 and was not restricted by language or year of publication to ensure the inclusion of all relevant studies. Further studies were retrieved from Google Scholar and JMIR and by hand searching reference lists.

Studies retrieved by the search strategy were exported to the web-based platform Covidence [55]. This allowed 2 authors (SB and AC) to independently review titles and abstracts—and then the full-text versions—against the inclusion and exclusion criteria via individual log-ins (Multimedia Appendix 2). The authorship and results of the studies were not masked. Any disagreements that arose were resolved through discussions between SB and AC.

Risk of Bias

The Downs and Black [56] (modified) checklist for randomized and nonrandomized studies was used to appraise study quality

[57]. Independently, 2 authors (SB and AF) rated 5 main assessment areas—the reporting, external validity, internal validity based on bias, internal validity based on cofounding and selection bias, and power—to provide an overall score out of 28. A score of 24 to 28 was graded excellent, 19 to 23 was graded good, 14 to 18 was graded fair, and <14 was graded poor [57]. Any disagreements between SB and AF in these ratings were resolved through discussion and re-examination of the study.

Summary Measures and Synthesis

To assist with data collection, a data spreadsheet was developed using Microsoft Excel to organize the data. Data collection included study characteristics, population, eHealth and mHealth interventions, outcome measurements, and study findings. Data collection was completed by 1 author (SB) and checked by all authors. A narrative synthesis method was used for methodological heterogeneity to identify and present common statistical descriptions [58]. All results were interpreted within the context of each study against the total number of studies and the considered risk of bias.

Where possible, data outcomes from similar studies were pooled, and a meta-analysis was performed to allow the comparison of an intervention group (IG) with a control group (CG). Baseline (time point 1) and end-of-study scores (time point 2) were entered into Review Manager software (RevMan version 5.4) to determine standardized mean differences (SMDs) and 95% CIs [59]. Forest plots were established using continuous data and a random-effects model because of the anticipated effect of clinical heterogeneity and to provide a summary of the distribution of effect [60]. A subanalysis was also conducted to reduce statistical heterogeneity. For the studies examining the same intervention and same fixed parameter, continuous data and a random-effects model were used to consider the common effect of the intervention [61].

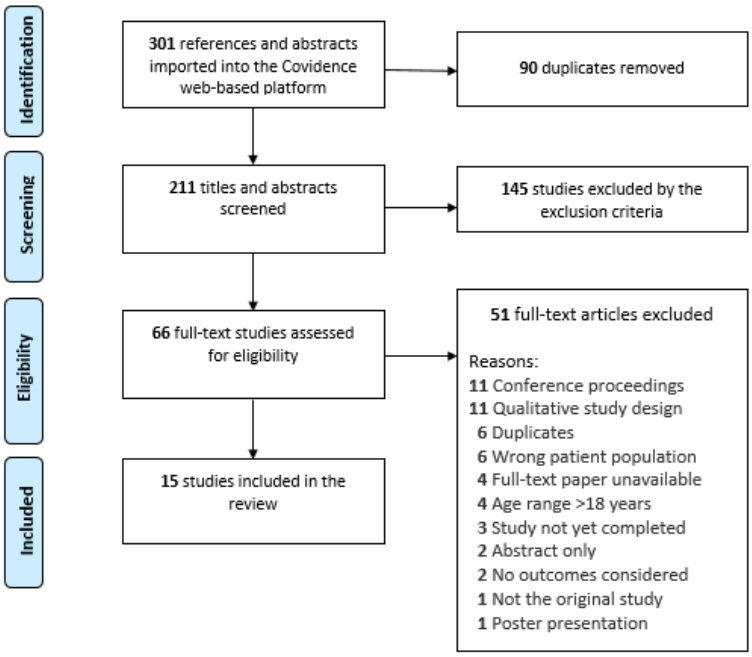
Finally, conclusions were drawn by visually inspecting forest plots and interpreting SMDs using the Hedge (adjusted) g. An effect size of 0.2 was considered small, 0.5 was considered medium, and 0.8 was considered large [62]. The presence of heterogeneity was also considered using I^2(I^2=100%×Q [chi-square]−df). A variation of 25% was reported as low, 50% was reported as moderate, and 75% was reported as high [63]. A P value of <.05 was considered statistically significant [64].

Results

Study Selection

The database search retrieved 301 studies. Of the 301 studies, 90 (29.9%) were duplicates; 145 (48.2%) did not meet the inclusion criteria based on their title or abstract; and 51 (16.9%) were excluded in the full-text screening because of study design, population, age range, outcomes, or the inability to gain the full text (eg, abstract only, conference presentations, and posters). Approximately 5% (15/301) of studies met the inclusion criteria to be introduced into this review (Figure 1) [65-79]. Of the 15 studies, only 1 (7%) was retrieved in a language other than English (Dutch), and an English version of the same study was attained through ResearchGate [72].

Figure 1. Summary of the study selection process using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.



Study Characteristics

Overview

The 15 studies included in this review were conducted in four countries: Canada [69,70,73-79], the Netherlands [65,67,68,72], the United States [66], and the United Kingdom [71]. These studies were published between 2008 and 2021 (Table 1) [65-79].

Table 1. Characteristics of the 15 studies showing population, intervention, control, outcomes, and study design.

First author and country	Population (N) and age range or mean (SD; years)	Intervention	Control (n)	Outcomes	Study design	Dropout (n)
Armbrust et al [65], the Netherlands	49; 8.7-10.8	Rheumates@Work	21[a]	Physical activity (effectiveness)	Multicenter observer blinded RCT[b]	7[c]
Connelly et al [66], United States	289; 12-18	Teens taking charge: managing arthritis on the web	144[d]	Self-management (effectiveness)	2-arm parallel group RCT	24[e]
Doeleman et al [67], the Netherlands	72; 10.6-16.4	EQ-5D-Y-5L[f] via Ruema2Go App	N/A[g]	HRQoL[h] monitoring to detect disease activity (efficacy)	Retrospective monocentric study	0[i]
Haverman et al [68], the Netherlands	176; mean 11.6 (SD 4.5)[j]	ePROfile	67	HRQoL (effectiveness) and PR[k] feedback (n=3)	Sequential cohort study	—[l]
Heale et al [69], Canada	31; 12.8-18.6	Wearable accelerometer using Misfit Flash	N/A	Physical activity (feasibility)	Pre- and postintervention design	3[e]
Lalloo et al [70], Canada	60; mean 15 (SD 1.7)[j]	iCanCope	29[d]	Self-management (feasibility and effectiveness)	2-arm pilot parallel group RCT	12[e]
Lee et al [71], United Kingdom	14; 7-16	My Pain Tracker	1 of 4 rotating groups[d]	Pain (effectiveness)	Randomized N-of-1 crossover trail	0
Lelieveld et al [72], the Netherlands	33; 8-12	Rheumates@Work	16	Physical activity (effectiveness)	Pilot RCT	0
Stinson et al [73], Canada	333; 12-18	Teens taking charge: managing arthritis on the web	169[d]	Self-management (effectiveness)	2-arm parallel group RCT	114[e]
Stinson et al [74], Canada	39; 12-17	iPeer2Peer Program	15[a]	Self-management (feasibility, usability, and effectiveness)	Pilot RCT	9[e]
Stinson et al [75], Canada	70[c]; age not available	eOuch	N/A	Pain (feasibility)	Correlational research	—
Stinson et al [76], Canada	101; 4-18	SUPER-KIDZ	N/A	Pain (efficiency) and PR feedback (n=15)	Descriptive design and 2-stage Delphi technique	—[m]
Stinson et al [77], Canada	46; 12-18	Teens taking charge: managing arthritis on the web	24	Self-management (feasibility)	Pilot RCT	9[c]
Stinson et al [78], Canada	13; 9-18	eOuch	N/A	Pain (feasibility and usability)	Descriptive study	3[n]
Stinson et al [79], Canada	112; 9-17	eOuch	N/A	Pain (feasibility and usability)	Prospective descriptive study	2

[a]Waitlist control.

[b]RCT: randomized controlled trial.

[c]Intention-to-treat analysis.

[d]Active control group.

[e]Excluded in final analysis.

[f]EQ-5D-Y-5L: EuroQol 5-dimensional youth 5-level.

[g]N/A: not applicable.

[h]HRQoL: health-related quality of life.

[i]Data from 4 children and young people were misinterpreted in the assessment and excluded from analysis.

[j]Age range not available.

[k]PR: pediatric rheumatologist.

[l]Not provided.

[m]Pain assessments were completed by parents instead of children (n=4; 4-7 years) and, therefore, excluded from the analysis.

[n]Dropouts (n=3) replaced in phase 2.

Participants

A total of 1438 children and young people (range 13-333) were included in this review [65-79]. Studies recruited children and young people from pediatric rheumatology centers or pediatric rheumatology departments within children's hospitals [65-70,72-79]; one of the studies recruited children and young people from the Childhood Arthritis Prospective Study [71].

Approximately 93% (14/15) of the studies reported on children and young people characteristics [65-74,76-79]. The mean age was 12.97 (SD 1.85) years, varying across studies between 9.7 years and 15.1 years. Most children and young people were female (887/1237, 71.7%) compared with males (350/1237, 28.29%), ranging from 62.9% to 96.7% [65-74,76-79]. The JIA subtypes were aligned with the International League for Rheumatology criteria [80]. Across the study population, the most common subtype of JIA that was reported was oligoarthritis, making up between 21% to 61% [65-74,76-79]. Approximately 7% (1/15) of the studies did not include these characteristics (N=70) [75], and 27% (4/15) of studies excluded the characteristics of children and young people lost during follow-up (n=123) [73,74] or excluded from the final analysis (n=8) [67,76]; (Multimedia Appendix 3 [65-79]).

Approximately 87% (13/15) of the studies considered disease activity [65-70,72-74,76-79]. Reporting either the mean range of disease (range 0.1 to 3.75) using the Physician Global Assessment, Juvenile Arthritis Disease Activity Score or 0-10cm Visual Analogue Scale [65,67-69,74,76-79], or the number of children and young people with low (range 60%-82.5%) scores, or moderate-to-high (range 17.4%-25%) scores [66,70,73], or the number of active (87%, 13/15) and inactive cases (13%, 2/15) [72].

Approximately 20% (3/15) of studies also included feedback from a range of pediatric rheumatology health care providers [67,68,76]. This included PRs (n=18; range 3-15) using eHealth interventions during consultations [68,76] or members of the Childhood Arthritis and Rheumatology Research Alliance (PRs and allied health) replying to a survey (survey 1:115 members; survey 2:157 members [73% replied to survey 1]) or attending a 2-day consensus conference (20 members; pediatric pain and rheumatology experts). Childhood Arthritis and Rheumatology Research Alliance members were from the United States and Canada [76].

Interventions

In total, 10 interventions were identified to support children and young people with JIA. The interventions were categorized according to their clinical aim to align with our research question, resulting in the formation of three themes: symptom monitoring, physical activity promotion, and self-management development.

Theme 1: Symptom Monitoring

Approximately 33% (5/15) of studies focused on self-reporting pain [71,75,76,78,79]. The interventions used included the following:

1. *My Pain Tracker*, an mHealth app aimed at monitoring pain 1 to 3 times a day or when needed [71]
2. *eOuch*, a customized electronic pain diary aimed at monitoring pain 3 times a day [75,78,79]
3. *SUPER-KIDZ*, a web-based assessment to self-report pain before consultations [76]

Approximately 13% (2/15) of studies focused on self-reporting HRQoL [67,68]. The used interventions included the following:

1. *EuroQol 5-dimensional youth 5-level questionnaire* (EQ-5D-Y-5L), accessed through the Reuma2Go health app aimed at remotely identifying disease activity and the need for treatment adjustments [67]
2. *ePROfile*, a web-based assessment (Kwaliteit van leven in kaart or quality of life map website) aimed at improving HRQoL discussion during rheumatology consultations [68]

Theme 2: Physical Activity Promotion

Approximately 20% (3/15) of studies focused on promoting physical activity [65,69,72]. The interventions used included the following:

1. A *wearable activity tracker*—using the commercially available *MisFit Flash*—aimed at improving physical activity levels (PALs) [69]
2. *Rheumates@Work*, a web-based behavioral and cognitive program aimed at delivering health information related to JIA and improving PALs [65,72]

Theme 3: Self-management Development

Approximately 33% (5/15) of studies aimed to develop self-management skills [66,70,73,74,77]. The interventions used included the following:

1. *Teens taking charge: managing arthritis online*, which is a web-based behavioral and cognitive program aimed at providing disease-specific information and self-management strategies to improve health outcomes [66,73,77]
2. *iCanCope*, a smartphone app aimed at tracking and improving pain self-management [70]
3. *iPeer2Peer Program*, a web-based peer-mentoring program aimed at facilitating positive role modeling and social support through video calls [74]

The expected level of engagement with the interventions varied from a few minutes before rheumatology consultations to 17 weeks [65-79]. Of the 15 studies, 14 (93%) required the children and young people to use the intervention at home (age range 7-18 years) [65-75,77-79], and only 1 (7%) was conducted in a clinical setting (age range 4-18 years) to monitor use [76]. For a more detailed description of each intervention, see Multimedia Appendix 4 [65-79,81].

Comparator or Control

Approximately 53% (8/15) of studies compared a pretested (time point 1) and posttested (time point 2) IG (455/904, 50.3%; range 17-144, median age 12.9, SD 2.09 years; female 322/455, 70.8%) with a CG (449/904, 49.7%; range 14-145, median age 13.4, SD 1.91 years; female 352/449, 78.4%) [65,66,68,70,72-74,77]. Of these 8 studies, 3 (38%) compared the IG with a CG receiving usual care (no eHealth or mHealth input) [68,72,77], 2 (25%) used a waitlist control method to allow all children and young people exposure to the intervention before study completion [65,74], and 3 (38%) compared the IG with an active CG also receiving a digital intervention [66,70,73].

One of the studies compared different real-time reporting schedules across 4 groups (n=12) with a median age of 12.5 years (range 10-14 years; female 9/12, 75%) [71].

Outcomes

Study outcomes varied according to the intervention stage of development (feasibility, usability, efficiency, and effectiveness) [65-79]. Health outcomes that considered an evaluation measurement to allow the quantitative comparison between groups, and an effectiveness analysis, were categorized to support the clinical aim of each intervention under the three intervention themes: symptom monitoring, physical activity promotion, and self-management development (Table 2).

Table 2. Formation of themes, evaluation criteria, and main outcomes supporting the delivery of the eHealth and mobile health interventions for juvenile idiopathic arthritis.

Theme (interventions aim)	Outcomes (evaluation measurement)
Theme 1: symptom monitoring	
Real-time pain	• Pain intensity, unpleasantness, interference using electronic VAS[a] 5 cm [75,78,79] and RPI[b] [75,78,79] • Pain location and descriptors: size (severity), throb rate (intensity), and emotion, PROMIS[c] and Pediatric pain Interference Scale–Short Form [71] • PedsQL[d] generic inventory—and arthritis module and PCQ[e] and Physician Rated Disease Activity Indices [79] • Children and young people aged 4-7 years: Faces Pain Scale–Revised; children and young people aged 8-18 years: NRS[f] (0-10 cm) [76], survey, and consensus conference [76]
HRQoL[g]	• EuroQol 5-dimensional youth 5-level questionnaire and 0-100 cm VAS (current health status) • Juvenile Arthritis Disease Activity Score with 71 joint count [67]
Pediatric rheumatology feedback	• HRQoL communication during pediatric rheumatology consultation, number of psychologist referrals, and PR[h] satisfaction [68] • Satisfaction questionnaire and 2-stage Delphi survey [76]
Theme 2: physical activity promotion	
Objective measurements	• Bruce Treadmill protocol for exercise capacity (endurance time) [65,72] • Accelerometer (Actical Phillips Respironics) for physical activity [65]
Self-reporting measurements	• A 7-day activity diary [65,72] • 3-Day Activity Recall to measure the metabolic equivalent values of activities and PROMIS [69] • PedsQoL (version 4) and pain and well-being (0-10 cm VAS); functional ability: CHAQ[i] [65,69] • School absenteeism, participation in physical education classes, and follow-up 3 and 12 months [65]
Functional capacity	• CHAQ [69] • Dutch version CHAQ38 [65]
Disease activity	• Disease and medication use records [72] • Disease activity 0-10 cm VAS [65] • Physician Global Assessment 0-10 cm or 0-100 cm VAS [69,72]
Theme 3: self-management development	
Pain reduction	• RPI [74,77] • Pain intensity [66,70,73] and interference [66,73] using an 11-point NRS (0-10) [66,70] • Tracking logs [70] • Follow-up at 3, 6, and 12 months [66,73]
HRQoL improvement	• PedsQL [66,70,73,74] • Juvenile Arthritis Quality of Life Questionnaire [77] • PROMIS: pediatric anxiety short form and depressive symptoms short form [66,73] • Perceived Stress Questionnaire [77] • Follow-up at 3, 6, and 12 months [66,73]
Functional capacity	• Child Activity Limitations Interview [70]
Health literacy	• Medical Issues, Exercise, Pain, and Social Support Questionnaire [66,73,74,77] • Children's Arthritis Self-Efficacy scale [66,73,74,77] • PCQ (behavioral and cognitive pain-coping strategies) [66,73] • Follow-up at 3, 6, and 12 months [66,73]
Adherence to prescribed treatment	• Child Adherence Report Questionnaire and Parent Adherence Report Questionnaire [73,77]

[a]VAS: visual analog scale.

[b]RPI: Recall Pain Inventory.

[c]PROMIS: Patient-Reported Outcomes Measurement Information System.

[d]PedsQL: Pediatric Quality of Life Inventory.

[e]PCQ: Pain Coping Questionnaire.

[f]NRS: numeric rating scale.

[g]HRQoL: health-related quality of life.

[h]PR: pediatric rheumatologist.

[i]CHAQ: Childhood Health Assessment Questionnaire.

Study Design

Study designs included two 2-arm parallel group randomized controlled trials (RCTs) [66,73], one 2-arm pilot parallel group RCT [70], 1 multisite observer-blinded RCT [65], 3 pilot RCTs [72,74,77], 1 randomized N-of-1 crossover trial [71], 1 descriptive study with 2-stage Delphi technique [76], 1 descriptive study with 2-phase testing [78], 1 prospective descriptive study [79], 1 retrospective monocentric study [67], 1 pre- and postdesign study [69], 1 correlational study [75], and 1 sequential cohort intervention study [68] (Table 1) [65-79].

Methodological Quality of Studies

Using the Downs and Black [56] (modified) checklist, the overall mean quality score of the 15 studies was 18.87 (SD 1.92) [65-79]. The scores ranged from 15 to 21, providing a fair-to-good score [57] (Multimedia Appendix 5 [65-79]). There were no disagreements between SB and AF that needed to be resolved by a third author (AC). The 2 areas in which study quality was consistently limited were power and sampling bias; 87% (13/15) of studies had insufficient power to detect a clinically significant effect [65,67-78], and convenience sampling and selection bias may have prevented full representation of the JIA population [68,72,75,79]. Children and young people were selected because of pain experience [66,70,78]; level of disease activity [65,66,69,71,75,78,79]; unlikelihood of medication changes [69]; no other comorbidities or cognitive impairments [65,66,69-71,73-75,77,79]; good visual acuity [75,79]; no hand deformities [75,79]; reduced PALs [65]; access to a computer, tablet, or phone or the internet [65,69,70,72,74]; and level of comprehension and ability to speak and read English [65,66,70,73-77,79], Dutch [72], Spanish [66], or French [73,77]. Methodological concerns were also seen in internal validity because of contamination or unreliable compliance [65-67,69-71,73,75-79].

Results of the Studies

Theme 1: Symptom Monitoring

Real-time Pain

Approximately 33% (5/15), fair-to-good–quality studies reported on real-time pain assessments [71,75,76,78,79]. Of these 5 studies, 3 (60%) reported on children and young people (aged 11.2-18 years) using *eOuch* to record their pain 3 times a day against the three pain rating measurements: intensity, unpleasantness, and interference [75,78,79], demonstrated a strong correlation (r=0.71-0.74, P<.01) between these pain measurements [79]. On average, pain scores reported were mild to moderate, interfering mostly with walking and least with school work, relationships with friends or family, and sleeping [78,79]. A good-quality study demonstrated changes in children's and young people's pain recordings throughout the day (interference P<.01, stiffness P<.01, and fatigue P<.01) and, week to week (intensity P<.01, unpleasantness P<.01, interference P<.01, and stiffness P<.01) [79]. Predicted changes in pain were also seen after a joint injection (medium effect size: 0.52-0.71); the main effect was for pain intensity [79]. A weak effect was reported for tiredness (r=0.24-0.26) and perceived ability to control pain (r=0.6-0.26) [79]; Multimedia Appendix 6 [71,75,78,79]).

Of the 5 studies, a further 1 (20%) fair-quality study reporting on the intervention *SUPER-KIDZ* that targeted children and young people aged between 4 to 18 years considered the pain dimensions that should be included in a pain assessment [76]. Using a 2-stage Delphi technique, the consensus view from health care experts (survey 1: n=115; survey 2: n=157; 2-day conference: n=20) concluded the inclusion of the characteristics of pain—intensity, location, frequency, duration, and the consequences of pain—and functional limitations [76].

Another 20% (1/5) of fair-quality studies reported on the frequency of recording real-time pain scores using *My Pain Tracker* [71]. Children and young people (aged 7-16 years) adherence rates were higher when pain was reported once a week (15/24, 63%) compared with when pain was reported once a day (85/168, 50.6%) or twice a day (127/336, 37.8%) or *as and when* pain was experienced (range 0-7 reports) [71]. There were no significant differences in pain interference scores because of reporting frequency (P=.77) or the different time points (weeks) across the study (r=−.004; P=.68). The children and young people qualitative results reported that they preferred once a day or *as and when* (6/14, 43%) reporting schedules [71] (Multimedia Appendix 6).

Real-time Pain Assessments Versus Recall Pain Assessments

Of the 5 real-time pain assessment studies, 3 (60%) fair-to-good–quality studies considered the correlation between *eOuch* real-time pain recordings and the Recall Pain Inventory short form [75,78,79]. For CPY (aged 11.2-18 years), a moderate to strong correlation (r=0.49-0.84) was reported between the real-time pain recordings and recall pain recordings (P<.01) [79], and the magnitude of changes in pain did not differ significantly when pain was defined as >0/100 or >0/30. However, when pain was defined as >0/10, there was weak *within-person consistency*, resulting in an 8% variance and a moderate association between the 2 assessments [75]. The same study also reported computed changes in pain (P=.02) against the judged assessment of pain (P=.004), finding both to be significantly similar, although the Recall Pain Inventory was higher and predictable [75]. Recall pain assessment measurements were mostly influenced by the children and young people peak pain score and the last real-time pain score. This finding appeared to be clinically significant (Multimedia Appendix 6).

Real-time Pain Scores Versus Other Commonly Used Pediatric Assessments

Of the 5 real-time pain studies, 1 (20%) fair-quality study compared real-time pain scores, using *eOuch*, with other pediatric tools (Pediatric Quality of Life Inventory [PedsQL]

Generic Inventory, PedsQL Arthritis Module, and Pain Coping Questionnaire). For children and young people (aged 9-17 years), a weak to moderate correlation (r=0.02-0.64) was seen, highlighting differences in the assessment tools, suggesting the need for specific pediatric pain assessments (Multimedia Appendix 6).

HRQoL Assessment Versus Disease Activity Assessment

Of the 15 studies, 1 (7%) good-quality study compared children and young people (aged 10.6-16.4 years) self-reporting HRQoL at home, using the *EQ-5D-Y-5L* assessment, with the commonly used clinical care tool Juvenile Arthritis Disease Activity Score with 71 joint count, which was completed by the PR during consultation to measure disease activity [67]. The HRQoL assessment (EQ-5D-Y-5L sum score) across all 5 levels (mobility, self-care, daily activities, pain or discomfort, and anxiety or depression) displayed satisfactory diagnostic accuracy (87%; 95% CI 76-94; P<.001), sensitivity (85%), specificity (89%), and predictive values (positive 88% and negative 86%) in identifying moderate to high disease activity [67]. This suggests that disease activity would not have been missed through remote monitoring of HRQoL, and treatment adjustments based on the current-to-treat guidelines (>1.5 for oligoarthritis and >2.5 for polyarthritis) could be applied [67].

PR's Feedback

Of the 15 studies, 1 (7%) good-quality study compared the preferred method of reviewing pain assessments by PRs (11/15, 73% female; 10/15, 67% practicing >10 years) [76]. *SUPERKIDZ* pain assessments were completed by children and young people (aged 4-18 years; with no help from parents) before the PR consultation using three different methods: a laptop or computer, a multimedia player, and a paper-based assessment. PRs (10/15, 67%) reported the electronic assessments to be more time efficient (P=.02) than the paper-based assessment and would recommend the use of web-based pain summaries to colleagues (9/15, 60%). There were no differences reported in developing pain management plans (10/15, 67%) [76].

Of the 15 studies, 1 (7%) fair-quality study reported on the PR's review of the web-based HRQoL assessment, *ePROfile*, during consultation [68]. PRs (n=5) reviewed 176 children and young people (mean 11.6, SD 4.5 years) tabulated answers and were satisfied with the care they provided for the IG compared with the care they provided for the CG, particularly in the areas of emotional support (first consultation [time point 1] P<.01 and second consultation [time point 2] P<.001) and meeting children and young people needs (time point 1 and time point 2 P<.001). PR satisfaction increased slightly in the second consultation compared with that of the first. PR evaluations reported *ePROfile* as useful (time point 1: 97/102, 95%; time point 2: 64/64, 100%), and the number of referrals increased (time point 1=9.2% and time point 2=4.3%) compared with the CG (3%). These results were not significant [68]. Parents also evaluated *ePROfile* as useful (time point 1: 57/65, 88%; time point 2: 37/46, 80%); however, parent satisfaction did not differ between the IG and CG, and children and young people (mean age 11.6, SD 4.5 years) reported the consultation as normal (time point 1: 47/48, 98%; time point 2 29/35, 83%). *ePROfile* was considered by the study authors as an efficient medium for monitoring HRQoL and was implemented in clinical use after the study [68].

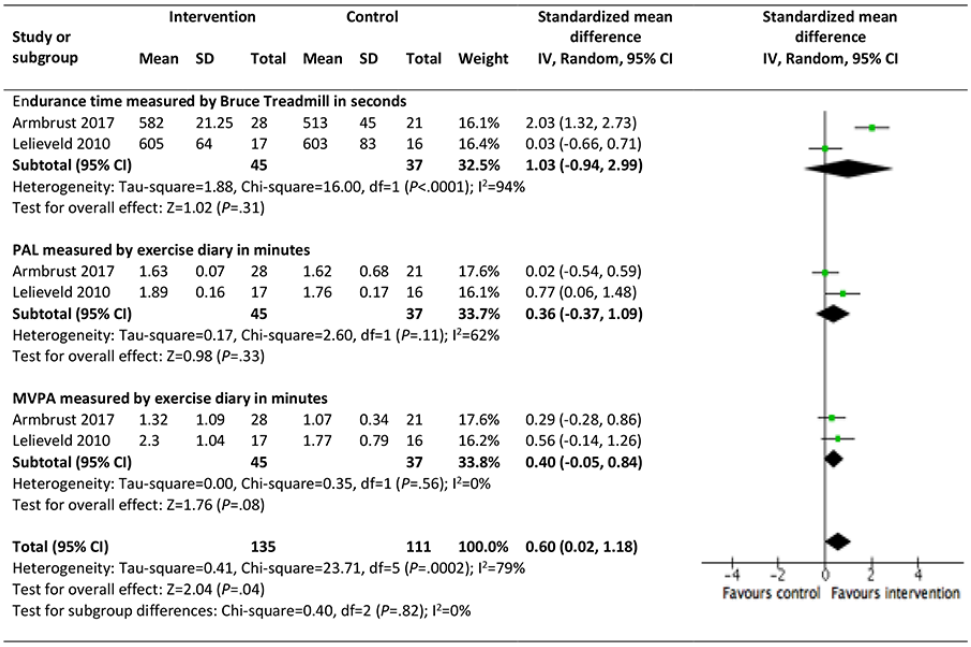
Theme 2: Physical Activity Promotion

Overview

Approximately 20% (3/15) of the fair-to-good–quality studies considered the interventions' effect on physical activity of children and young people [65,69,72]. Of these 3 studies, 1 (33%) fair-quality study reported on children and young people (aged 12.8-18.6 years) wearing an *activity tracker*, Misfit Flash, daily for 28 days. No significant differences in PALs were recorded [69].

The other 67% (2/3) good-quality studies reporting on children and young people (aged 8.7-12 years) who used the intervention *Rheumates@Work* were pooled in a meta-analysis [65,72]. Overall, a moderate effect (SMD 0.60, 95% CI 0.02-1.18, P=.40) was seen in physical activity (endurance time, PAL, and moderate to vigorous physical activity [MVPA]). However, there was high statistical heterogeneity between the studies (I^2=79%), suggesting a 79% variance across the studies, reducing confidence in these results (Figure 2) [65,72]. No changes were reported for pain intensity, disease activity, or functional ability [65,69].

Figure 2. Effectiveness of Rheumates@Work on the promotion of physical activity for juvenile idiopathic arthritis (aged 8-12 years). MVPA: moderate-to-vigorous physical activity; PAL: physical activity level.



Seasonal Intervention Effect

Of the 3 studies considering physical activity promotion, 1 (33%) good-quality study, *Rheumates@Work*, reported a seasonal intervention effect after comparing a winter IG to a summer IG. For the winter IG, a 24-minute reduction in rest was recorded using an accelerometer (Actical Phillips Respironics). This result was significant (*P*=.05) [65].

Follow-up

Of the 3 studies considering physical activity promotion, only 1 (33%) good-quality study, *Rheumates@Work*, considered follow-up after the study period [65]. At 3 months, for the IG, children's and young peoples' (aged 8.7-18 years) physical activity (endurance time and PAL) continued to improve, and by 12 months, it declined. However, this reduction did not reach the preintervention levels. Positive improvements were also reported for educational participation. At 3 months, school absenteeism decreased from 43% to 14% (*P*=.02) in the IG and increased from 24% to 29% (*P*=.60) in the CG. Children's and young peoples' participation in physical education classes also improved in the IG group, from 57% to 71% (*P*<.01) and from 62% to 67% in the CG (*P*=.01). However, these differences were not statistically significant [65].

Theme 3: Self-management Development

Overview

Approximately 33% (5/15) of fair-to-good quality studies assessed the health-related benefits of self-management development [66,70,73,74,77].

Pain Reduction

Of the 5 studies promoting self-management, all (100%) fair-to-good–quality studies monitored for changes in pain because of the intervention [66,70,73,74,77]. Of these 5 studies, 1 (20%) fair-quality study reported on children and young people (mean age 12, SD 1.7 years) using *iCanCope* [70]. The IG received a pain monitoring and self-management program, and the CG received pain monitoring only. Both groups reported a reduction in pain intensity (IG: 1.73-point reduction; CG: 1.09-point reduction), using a 0 to 10 numerical rating scale. These results were not statistically significant (*P*=.24) [70].

Of the 5 studies, 4 (80%) good-quality studies (children and young people aged 8-18 years) reporting on *Teens taking charge* and the *iPeer2Peer Program* were pooled for a meta-analysis [66,73,74,77]. A small postintervention effect was seen in the IG compared with the CG in reducing pain intensity (SMD −0.14, 95% CI −0.43 to 0.15; I^2=53%; *P*=.33). However, these results were not statistically significant, and moderate statistical heterogeneity was seen between the studies (Figure 3) [66,73,74,77]. No effect was seen on pain interference (Figure 4) [66,73].

Figure 3. Effectiveness of self-management programs in reducing pain intensity for children and young people (aged 12-18 years) with juvenile idiopathic arthritis.

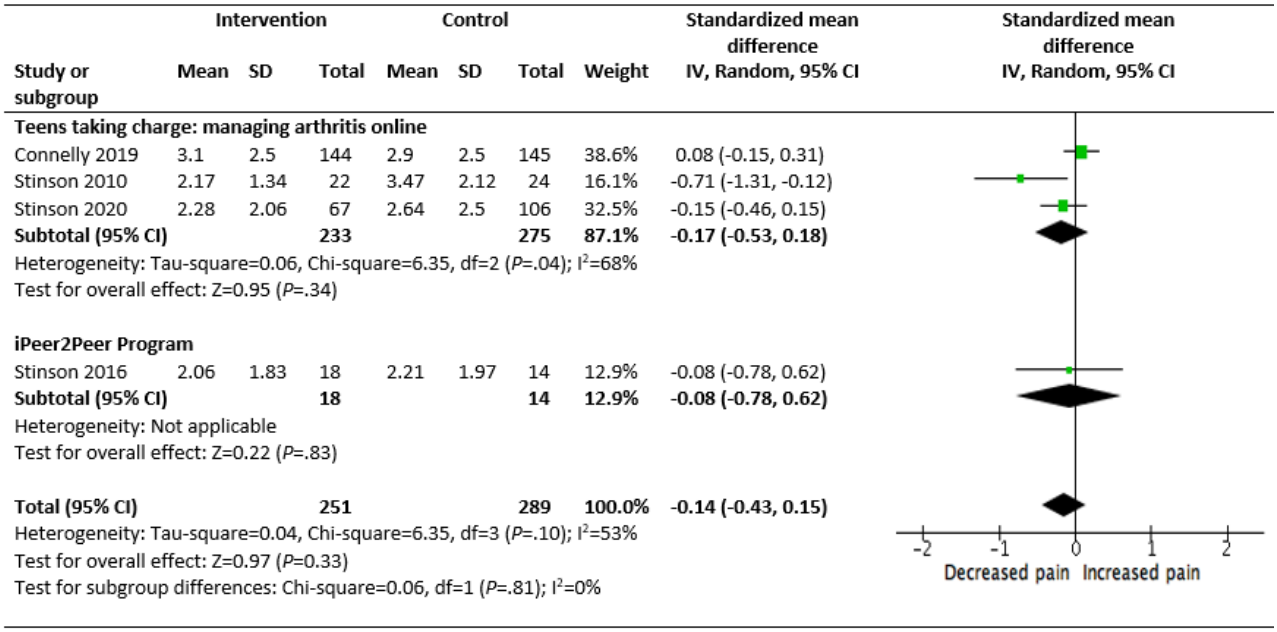
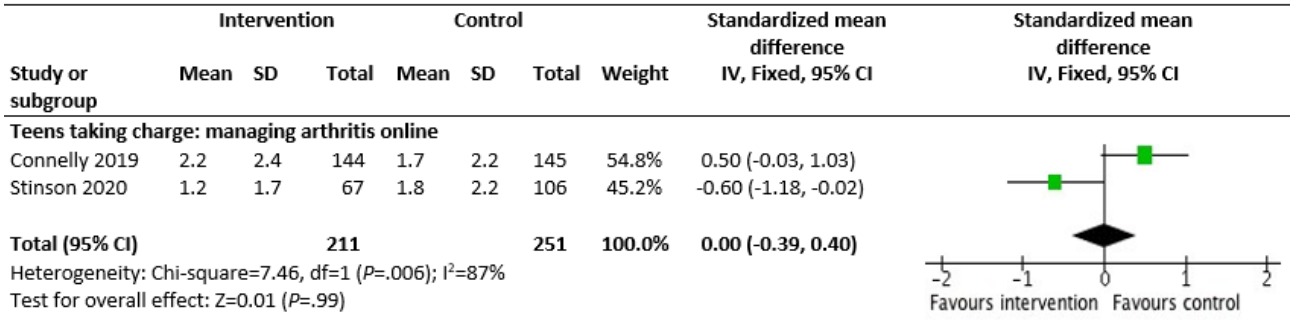


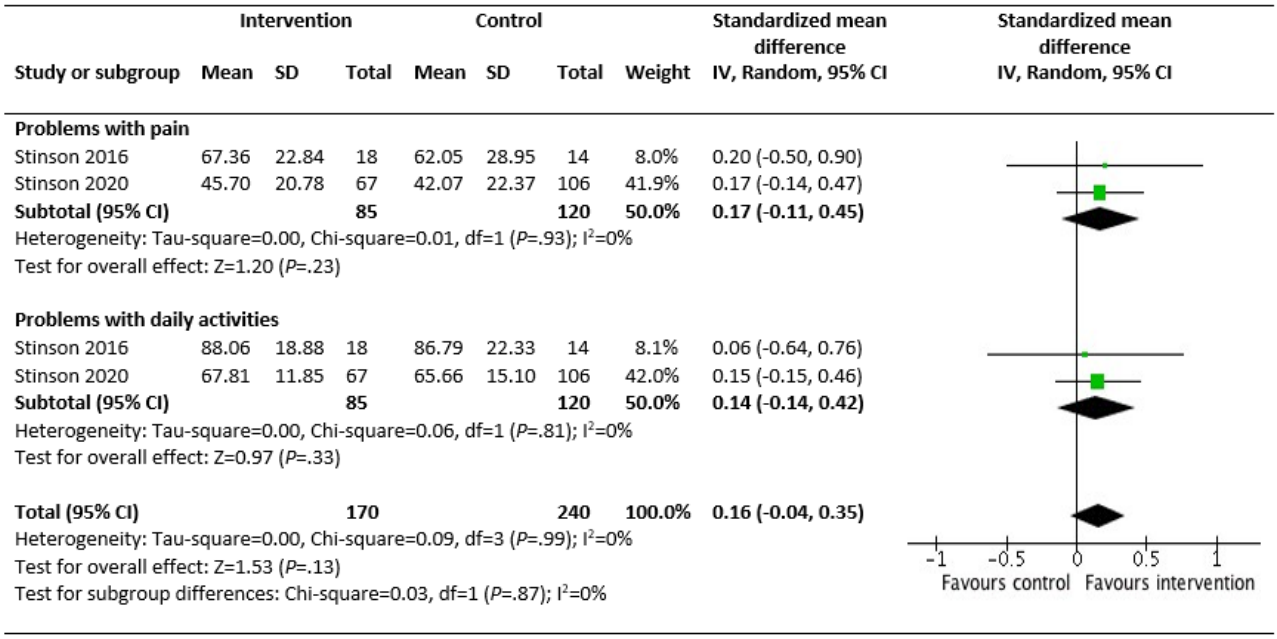
Figure 4. Effectiveness of Teens taking charge intervention in reducing pain interference for children and young people (aged 12-18 years) with juvenile idiopathic arthritis.



HRQoL Improvements

Of the 5 studies targeting self-management development, 4 (80%) fair-to-good–quality studies considered the intervention effect on HRQoL for the IG compared with CG (age range 8.7-18.6 years) [66,70,73,74,77]. Of these 5 studies, 4 (80%) good-quality studies, reporting on *Teens taking charge* and the *iPeer2Peer Program*, were pooled for a meta-analysis. No effect was demonstrated for HRQoL [66,73,74,77]. For *Teens taking charge*, a further subanalysis of the individual HRQoL domains (problems with pain, daily activities, treatment, worry, and communication), using the PedsQL, demonstrated a small effect in improving problems with pain and problems with daily activity (SMD 0.16, 95% CI −0.04 to 0.35; I=0%; *P*=.13) (Figure 5) [73,74]. This effect was not statistically significant. From the study outcomes excluded from the meta-analysis, no improvements were seen in anxiety, depression [66,73], or stress [77].

Figure 5. Subanalysis of Teens taking charge intervention and the health-related quality of life domains: problems with pain and daily activities for children and young people (aged 8.7-18 years) with juvenile idiopathic arthritis.



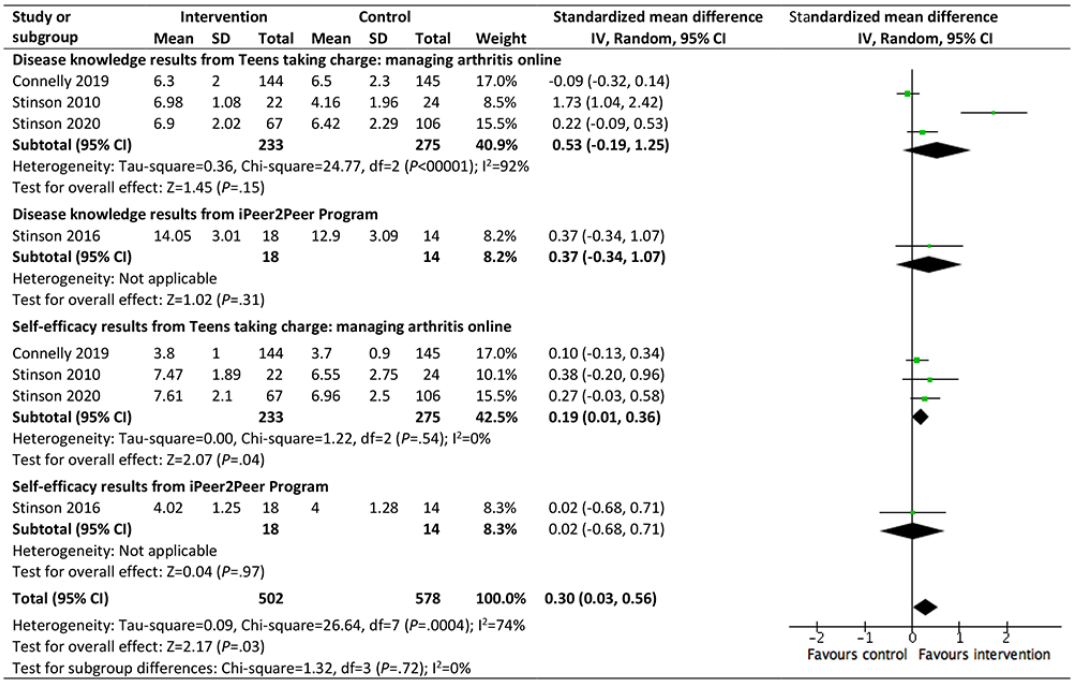
Follow-up

Of the 5 studies targeting self-management development, 2 (40%) good-quality studies considered follow-up after the study period at 3, 6, and 12 months [66,73]. In the Canadian *Teens taking charge* study, children and young people (aged 12-18 years) in the IG retained the improvements they gained during the study period for pain intensity and in the HRQoL domains of problems with pain and problems with daily activities. These results were not statistically significant. A significant improvement was seen in the domain of problems with treatment (*P*=.008) [73]. In the US *Teens taking charge* study, children and young people (aged 12-18 years) in the IG and CG continued to have a stable reduction in pain intensity and pain interference and improvements in HRQoL. The differences between the IG and CG were not significant [66].

Health Literacy

Of the 5 studies targeting self-management development, 4 (80%) good-quality studies, reporting on *Teens taking charge* and the *iPeer2Peer Program* (children and young people aged 8.7-18 years) and considering health literacy, were pooled in a meta-analysis [66,73,74,77]. A small, nonsignificant effect was seen in improving disease knowledge and self-efficacy (SMD 0.30, 95% CI 0.03-0.56; I^2=74%; *P*=.03); however, confidence in these results was reduced because of high statistical heterogeneity (Figure 6) [66,73,74,77]. No improvements were seen in pain coping strategies [66,73].

Figure 6. Effectiveness of self-management programs in improving disease knowledge and self-efficacy for children and young people (aged 12-18 years) with juvenile idiopathic arthritis.



Functional Ability and Adherence to Treatment

Of the 5 studies targeting self-management development, only 1 (20%) fair-quality study reported on functional ability. There was no improvement in pain-related limitations (*P*=.65) [70]. Another 40% (2/5) of good-quality studies reported on treatment adherence. No improvements were reported for medication, exercise, or splint adherence [73,77].

Adverse Events

Of the 15 studies, 3 (20%) fair-to-good–quality studies recorded adverse events [65,66,69]. *Teens taking charge* (age range 12-18 years) reported the highest number of adverse events (n=72), mostly related to infections (18/72, 25%) and arthritis-related flares (17/72, 24%) [66]. The more serious events involved hospitalization (9/72, 13%) or suicidal thoughts (4/72, 6%). There was no significant difference in adverse events between the IG and CG groups (*P*=.67) [66]. *MisFit Flash* (age range 12.8-18.6 years) also reported illness, injury, or pain (9/28, 32%), including arthritis-related ankle and knee pain (1/28, 4%). However, no significant difference was seen in functionality (mean Childhood Health Assessment Questionnaire score), pain, or active joint count during the study [69]. Whereas *Rheumates@Work* (age range 8-12 years) reported arthritis-related flares, affecting more children and young people in the CG (2/17, 12%) compared with the IG (1/16, 6%) [65].

Dropout

Of the 15 studies, 10 (67%) studies reported dropout rates (range 0-114) by children and young people (aged 8.7-18.6 years) [65,66,69,70,72-74,77-79]. Dropout reasons before study commencement included not being interested anymore [65,66,74], early withdrawal before allocation [66,77], not receiving allocation [65,70,77], not completing app orientation [70], technical issues [77], and no show and no reason [65]. Reasons during the study period included other health problems [65,69], school and extracurricular activities [69,78], discontinued use [66,74,77], did not complete final web-based measures [77], unable to reach [66], lost to follow-up [66,69,73,74,77,79], and removal because of lack of compliance [74]. No comparisons were made between age or gender [65,66,69,70,72-74,77-79].

Of the 15 studies, 7 (47%) studies reported both the IG and CG dropout rates [65,66,70,72-74,77]. A higher dropout rate was reported in the IG (119/455, 26.2%; range 0-76) compared with the CG (56/449, 12.5%; range 0-56) [65,66,70,72-74,77]. The Canadian *Teens taking charge* study reported the highest dropout rate (IG: 76/164, 46.3%; CG: 38/169, 22.5%) [73].

Discussion

Principal Findings

To the best of our knowledge, this is the first systematic review to evaluate the effectiveness of eHealth and mHealth interventions in supporting children and young people living with JIA. In total, 10 interventions were identified to support symptom monitoring, physical activity promotion, or self-management development for children and young people aged 4 to 18.6 years. These 10 interventions included 4 (40%) web-based programs [65,66,68,72,73,76,77], 3 (30%) health applications [67,70,71], 1 (10%) telecommunication application [74], 1 electronic diary (10%) [75,78,79], and 1 (10%) accelerometer compatible with a tablet or smartphone [69]. The methodological quality of the studies supporting these interventions ranged from fair [68-71,75,76,78] to good [65-67,72-74,77,79].

Theme 1: Symptom Monitoring (4-18 Years)

Pain assessment was the most common type of intervention used to support symptom monitoring. The interventions *My*

Pain Tracker and *eOuch* aimed to capture real-time data through children and young people self-reporting pain [71,75,78,79]. Monitoring pain is important as pain is the most frequently reported symptom by children and young people living with JIA [12,82]. Pain can dramatically interfere with physical functioning, coping mechanisms, and quality of life [12]. Stinson et al [79], through the use of *eOuch* pain diaries, demonstrated a correlation between pain intensity and the impact pain can have on emotional well-being (unpleasantness) and activities of daily living (interference), reinforcing the need for ongoing comprehensive pain monitoring, which could allow the health care team to make timely recommendations and prevent poor health outcomes [83-85].

However, there is no consensus on the required number of real-time assessments, per day or week [71,86,87], to ensure the collection of high-quality data and avoid the burden of momentary reporting [71]. Instead, a large variation, ranging from 2 to 9 times a day for children and young people, has been seen [87]. In this review, real-time pain monitoring ranged from 1 to 3 times a day [71,75,78,79] or when needed [71]. Lee et al [71], through *My Pain Tracker*, compared these reporting frequencies, finding that children and young people preferred once-a-week or when-needed pain assessments to avoid thinking about their pain. Although more details in pain data were collected from once-a-day reporting, and, for some children and young people, adherence to once-a-day reporting was easy as it became a routine [71], more research is needed on reporting frequencies.

Real-time pain monitoring also exposed differences between real-time pain and recall pain assessments [75]. Recall pain measurements were higher and predictive compared with average real-time pain measurements, influenced by the children's and young peoples' most intense pain and last pain score [75]. This is known as recall bias or peak-end effect [75,88]. This nonequivalence between real-time pain assessments and recall pain assessments adds significance to previous research by Stone et al [88], highlighting methodological concerns around relying on retrospective pain assessments, especially when considering the length between rheumatology appointments.

Longitudinal variances were also seen between real-time and recall pain monitoring [75]. Stinson et al [75] and Stone et al [88] both demonstrated a weak correlation with within-person data when pain was defined as >0/10, which is the most common pediatric pain scale. This suggests that real-time and recall pain assessments cannot be compared or used interchangeably when assessing long-term changes in pediatric pain [75]. Considering that the length of the studies in this review was only 2 to 8 weeks [71,75,78,79], further research on the longitudinal effects of real-time pain monitoring is needed.

The use of real-time symptom monitoring for children and young people is also supported by previous work in reducing the recall time to days, hours, or minutes [87], and importantly, a recent systematic review, reporting on real-time monitoring using mobile technology, suggests that it can be successfully implemented from the age of 7 years [87]. In addition, a study considering adults with chronic illnesses supports real-time monitoring for the identification of exacerbations, confidence in self-management, and prevention of hospital admissions [89]. In this review, the intervention *EQ-5D-Y-5L* endorsed this finding, as remote HRQoL monitoring identified, with satisfactory diagnostic accuracy (P<.001), moderate to high levels of disease activity, promoting the need for adjustments with prescribed treatments and rheumatology consultation frequency [67]. Further research is now needed on this web-based PR– children and young people interaction and the impact remote monitoring may have on safety [67,90].

In this review, 13% (2/15) of studies reported positive feedback from PRs after they reviewed web-based assessments during consultation [68,76]. PRs reported that *SUPER-KIDZ* pain assessments increased time efficiency compared with a paper-based assessment [76]. However, *ePROfile* increased PR satisfaction with the care they provided as the HRQoL discussion improved and the number of psychological referrals increased [68]. Although these findings were not significant, reviewing pain and HRQoL during consultation is important as children and young people with JIA have significantly lower HRQoL compared with that of healthy children and young people, and children aged 8 to 12 years with JIA have lower HRQoL than that of children with other chronic health conditions [9].

Interestingly, the use of web-based portals in adult rheumatology has been long standing. The Feed Forward System, for example, used in Sweden generates a patient's progress over a period and has been successfully used to guide health care provider recommendations and aid the development of patient self-management skills [91].

For JIA, feasibility studies considering web-based portals also support their use, reporting that this form of technology can increase children's and young peoples' (aged 5-22 years) feeling of control [92,93].

Regrettably, parents and children and young people did not report the same level of satisfaction with the *ePROfile* consultation as PRs [68]. Haverman et al [68] suggest that this may be as they are already happy with the quality of their care. Nonetheless, many factors that can influence children's and young peoples' opinions on digital assessments need to be considered. First, they can be influenced by the assessment experience; they need graphical and tailored feedback to encapsulate their results and catch their interest [94]. In addition, children and young people may not value and understand the importance of monitoring symptoms, disease, and general well-being (mood, fatigue, and functional ability) [44,95] or the need for a person-centered framework that builds partnerships between families and health care teams [92]. Further research on the use of web-based portals for children and young people is needed.

Theme 2: Physical Activity Promotion (8-18.6 Years)

In this review, 20% (2/10) of interventions, *Rheumates@Work* and the wearable *activity tracker*, Misfit Flash, aimed at improving self-management behavior by promoting physical activity for children and young people (aged 8-18.6 years) [65,69,72]. Of these 2 interventions, only 1 (50%)

Rheumates@Work, demonstrated a moderate but clinically meaningful effect on physical activity, improving endurance time, PAL, and MVPA for children and young people (aged 8-12 years) [65,72]. This finding is important as children and young people with JIA are less physically active [96,97] and spend more time in sedentary activities than their peers [96]. Improving physical activities helps to retain musculoskeletal function, muscle strength, and functional capacity [98].

In addition, increased physical activity did not exacerbate disease activity or pain in the IG compared with the CG [65,69]. In fact, no significant difference was reported [65,69], and absenteeism from school decreased [65]. These findings are encouraging, especially considering the related impact JIA can have on reducing academic performance, as depicted by Bouaddi et al [99] and Laila et al [100]. Although these findings are limited, they will add to the growing body of evidence reporting that exercise therapy is well-tolerated by children and young people with JIA [98,101,102], further supporting physical activity as a helpful and necessary treatment modality, improving adherence [24,103].

Theme 3: Self-management Development (8-18 Years)

In this review, 30% (3/10) of interventions—*iCanCope* [70], *Teens taking charge* [66,73,77], and the *iPeer2Peer program* [74]—supported self-management development for children and young people (aged 8-18 years). These interventions (including *Rheumates@Work* [65]) are typical behavior change technique interventions used for children and young people [31]. They support self-management through the development of disease-specific knowledge, goal setting, self-management strategies, and social support [66,70,73,74,77].

In this review, identified in the meta-analysis, children and young people participating in the self-management programs *Teens taking charge* and *iPeer2Peer Program* reported a small but nonsignificant improvement in pain intensity, disease knowledge, and self-efficacy scores [66,73,77]. However, high statistical heterogeneity was also seen within the results. This may be because of several reasons. First, a range of comparators was used for the CG. For example, in 2 of the 3 *Teens taking charge* studies, the IG was compared with an active CG rather than usual care. The CG also received an eHealth intervention 12 publicly available health education websites with phone support to support care. Improvements were then seen in both the IG and CG, reducing the mean difference between the groups [66]. Most digital studies primarily focus on a single intervention to demonstrate the intervention effect rather than comparing different digital interventions. However, it is this direct comparison that can reveal a more effective intervention [104]. The CG's intervention, the use of health care workers signposting quality health education websites to support self-management skills [105,106] and improve well-being [107], is supported in the literature. Therefore, acknowledging the improvements seen in the CG is important, as the use of publicly available websites can be a cost-effective solution for dissipating health information among the masses to support the delivery of health care [28]. For example, a study of adults living with chronic pain (n=20; aged 18-74 years) explained that if they had been provided with quality pain-related information, it might have prevented the desperation and anxiety they experienced, especially during the first few years [107].

Another explanation for the moderate to high statistical heterogeneity may have been that the studies were conducted in different countries, within different health care delivery systems, with different levels of pre-existing support [73]. Differences exist with the pediatric rheumatology workforce worldwide [7,13-18] and within publicly funded and self-funded health care systems [108,109]. Differences also exist in PRs' opinions on pediatric self-management and the use of an interdisciplinary approach to care [108]. Successful publicly available digital interventions may be a solution to transcending these boundaries and universally improving access to care [28,106]. Further comparisons between different self-management interventions is needed, especially when considering the dropout rates in this review, which, for the self-management programs, were higher in the IG [65,66,70,73,74]. These dropout patterns were similar to a recent systematic review predicting dropout rates in adults, with dropouts occurring at the beginning and over the course of the intervention [110].

There may also be no one-size-fits-all intervention, or there may be a need for a combination of interventions. For example, the *iCanCope* pain self-management application combined 2 interventions (real-time pain monitoring and self-management) and then compared this combination to standalone pain monitoring. This combination demonstrated a greater decrease in pain intensity scores (>1 point, 0-10 on the numerical rating scale) [70]. Although this finding was not significant, the inclusion of self-management programs could be clinically beneficial. Improving and providing effective educational interventions early in childhood should be when children and young people are beginning to develop their health behaviors [23]. Studies have shown that a high level of health literacy can support informed decision-making [111-113]; treatment adherence, especially for nonmedication interventions [113]; and the prevention of chronic health-related problems [22].

Unfortunately, not all the results of this review are promising. Across the studies, the interventions had no effect on pain interference [66,73], HRQoL [65,66,69,70,73,74,77], anxiety, depression, pain coping [66,73], disease activity [69,72], functional ability [65,69,70], or treatment adherence [73,77]. In addition, only 20% (3/15) of studies considered long-term follow-up [65,66,73]. More research is needed to gain wider health-related benefits.

Limitations

It is essential that several limitations are considered when interpreting the findings of this review. First, our search strategy was restricted to an academic context, using eHealth electronically indexed health databases that publish peer-reviewed journals, rather than apps within commercial stores. This means that our results may not provide a true reflection of the health apps available for JIA. This decision was based on the commonly reported shortcomings of health apps available to the general public that are related to data safety and lack of rigorous testing [114].

Second, the selection criteria in this review deviated from our systematic review protocol [53]. In the protocol, we outlined that the comparator or CG was to receive usual care, with no eHealth or mHealth input. Instead, we included 13% (2/15) of studies comparing an eHealth intervention to another digital intervention [66,73], as a preliminary pilot study of this intervention met our inclusion criteria to be included in this review [77]. This decision enabled us to provide the most up-to-date evidence for this intervention.

Third, our findings in this review supporting the use of real-time monitoring and web-based assessments were based on descriptive summaries. The use of a narrative, descriptive methodology to summarize, synthesize, and report the results is at risk of reporting bias. To reduce this risk, all authors internally reviewed all the stages of this review.

In addition, there were methodological concerns in the data reported by some studies because of performance bias. It was not possible to blind children and young people from the intervention, which could have resulted in a placebo effect. For example, *Rheumates@Work* reported improvements in both the IG and CG for MVPA and participation in physical education classes. Baseline testing may have made the CG more aware of the need to improve their physical activities [65]. The interventions' true effects may have also been overestimated, as activity levels entered by children and young people in the exercise diaries did not match the accelerometers. Overreporting exercise is not uncommon. Various correlations have been reported between exercise diaries and accelerometer recordings in the general population (r=0.52) [115]; among children and adolescents (reliable coefficient ranges r=0.5-0.93 and validity coefficient ranges r=0.03-0.88), with children being in the lower range [116]; and for JIA (light PAL and MVPA r<0.24; rest and PAL r=0.41) [117]. Unexpectedly, a 4% inaccuracy has also been identified in accelerometer recordings for JIA in light PALs (effect size 1.2) because of nonwearable periods (aquatic activities and ball games) [117]. Awareness of these possible variations enables correction. For example, Armbrust et al [117] recommend, for research purposes, the use of accelerometer recordings (7-19 days) and an activity diary (>13 days). Another feasible suggestion may be the use of wearable forms of digital technology (ie, a smartwatch) [27]; however, more research is needed to overcome the nonwearable periods such as contact sports [118] or while attending school [119].

Finally, the generalizability of our findings may be limited. We included 40% (6/15) of studies where several children and young people were categorized as unknown, not yet diagnosed, or other (55/1438, 3.82%; range 2-37) [68,70,73,76,77,79]. Dissecting the results to target children and young people specifically with JIA was not possible (Table 2). For a fair-quality study, which reported the highest number of children and young people in this category (37/101, 36.6%), our data extraction focused on the consensus view of pediatric rheumatology providers (PRs, allied health experts, and pain experts) rather than the children and young people [76], which is an area of research that is currently limited.

Conclusions

Evidence that supports the inclusion of eHealth and mHealth interventions in JIA management is on the rise; however, this evidence needs to be considered cautiously. Confidence in the results is reduced because of low sample size, wide CIs, high statistical heterogeneity, and no similar effect being seen across similar studies. More rigorous research is needed that focuses on the longitudinal effects of real-time monitoring, web-based PR–children and young people interactions, comparison of self-management strategies, and the use of wearable digital technology as an objective measurement for monitoring physical activity before any recommendations informing current practice can be given.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search terms and search strategy.
[DOCX File , 14 KB - jmir_v24i2e30457_app1.docx]

Multimedia Appendix 2
Inclusion and exclusion criteria.
[DOCX File , 15 KB - jmir_v24i2e30457_app2.docx]

Multimedia Appendix 3
Juvenile idiopathic arthritis subtypes based on the International League for Rheumatology criteria.
[DOCX File , 19 KB - jmir_v24i2e30457_app3.docx]

Multimedia Appendix 4
Overview of the eHealth and mobile health interventions used for juvenile idiopathic arthritis.
[DOCX File , 23 KB - jmir_v24i2e30457_app4.docx]

Multimedia Appendix 5

Methodological scores of the 15 studies using the Downs and Black [56] (modified) checklist.
[DOCX File , 18 KB - jmir_v24i2e30457_app5.docx]

Multimedia Appendix 6
Real-time pain monitoring versus other tools commonly used in pediatric rheumatology or pediatrics.
[DOCX File , 20 KB - jmir_v24i2e30457_app6.docx]

References

1. Palman J, Shoop-Worrall S, Hyrich K, McDonagh JE. Update on the epidemiology, risk factors and disease outcomes of Juvenile idiopathic arthritis. Best Pract Res Clin Rheumatol 2018 Apr;32(2):206-222. [doi: 10.1016/j.berh.2018.10.004] [Medline: 30527427]
2. Model of care for the NSW paediatric rheumatology network: muscoskelatal network. NSW Agency for Clinical Innovation. 2013. URL: https://www.aci.health.nsw.gov.au/__data/assets/pdf_file/0011/183656/HS13-027_ACI_PaedRheum_web.pdf [accessed 2022-01-14]
3. Ratcliff C, Denman G, Jandial S, Foster H. Diagnosing arthritis in children. Paediatrics and Child Health 2019 Dec;29(12):503-514 [FREE Full text] [doi: 10.1016/j.paed.2019.09.007]
4. Moorthy LN, Peterson MG, Hassett AL, Lehman TJ. Burden of childhood-onset arthritis. Pediatr Rheumatol Online J 2010 Jul 08;8:20 [FREE Full text] [doi: 10.1186/1546-0096-8-20] [Medline: 20615240]
5. Shoop-Worrall SJW, Kearsley-Fleet L, Thomson W, Verstappen SMM, Hyrich KL. How common is remission in juvenile idiopathic arthritis: A systematic review. Semin Arthritis Rheum 2017 Dec;47(3):331-337 [FREE Full text] [doi: 10.1016/j.semarthrit.2017.05.007] [Medline: 28625712]
6. Fingerhutova S, Saifridova M, Vranova M, Dolezalova P, Sebkova S, Bohm M, et al. Is there an evidence for the role of multidisciplinary team in the management of active juvenile idiopathic arthritis? Pediatr Rheumatol 2014 Sep 17;12(S1):177. [doi: 10.1186/1546-0096-12-S1-P177]
7. Spencer C. Why should pediatric rheumatology be recognized as a separate subspecialty: an open letter to medical councils and government agencies. Pediatr Rheumatol Online J 2007 Nov 21;5:21 [FREE Full text] [doi: 10.1186/1546-0096-5-21] [Medline: 18031581]
8. McErlane F, Carrasco R, Kearsley-Fleet L, Baildam EM, Wedderburn LR, Foster HE, et al. Growth patterns in early juvenile idiopathic arthritis: results from the Childhood Arthritis Prospective Study (CAPS). Semin Arthritis Rheum 2018 Aug;48(1):53-60 [FREE Full text] [doi: 10.1016/j.semarthrit.2017.11.002] [Medline: 29217290]
9. Haverman L, Grootenhuis MA, van den Berg JM, van Veenendaal M, Dolman KM, Swart JF, et al. Predictors of health-related quality of life in children and adolescents with juvenile idiopathic arthritis: results from a Web-based survey. Arthritis Care Res (Hoboken) 2012 May;64(5):694-703 [FREE Full text] [doi: 10.1002/acr.21609] [Medline: 22238240]
10. Munro J, Murray K, Boros C, Chaitow J, Allen RC, Akikusa J, Australian Paediatric Rheumatology Group. Australian Paediatric Rheumatology Group standards of care for the management of juvenile idiopathic arthritis. J Paediatr Child Health 2014 Sep;50(9):663-666. [doi: 10.1111/jpc.12462] [Medline: 25156704]
11. Quartier P. Current treatments for juvenile idiopathic arthritis. Joint Bone Spine 2010 Dec;77(6):511-516. [doi: 10.1016/j.jbspin.2010.09.002] [Medline: 20961790]
12. Stinson JN, Luca NJ, Jibb LA. Assessment and management of pain in juvenile idiopathic arthritis. Pain Res Manag 2012;17(6):391-396 [FREE Full text] [doi: 10.1155/2012/237258] [Medline: 23248812]
13. Foster H, Rapley T. Access to pediatric rheumatology care -- a major challenge to improving outcome in juvenile idiopathic arthritis. J Rheumatol 2010 Nov;37(11):2199-2202. [doi: 10.3899/jrheum.100910] [Medline: 21041261]
14. Al Maini M, Adelowo F, Al Saleh J, Al Weshahi Y, Burmester G, Cutolo M, et al. The global challenges and opportunities in the practice of rheumatology: white paper by the World Forum on Rheumatic and Musculoskeletal Diseases. Clin Rheumatol 2015 May;34(5):819-829 [FREE Full text] [doi: 10.1007/s10067-014-2841-6] [Medline: 25501633]
15. Cox A, Piper S, Singh-Grewal D. Pediatric rheumatology consultant workforce in Australia and New Zealand: the current state of play and challenges for the future. Int J Rheum Dis 2017 May;20(5):647-653. [doi: 10.1111/1756-185X.12802] [Medline: 26864133]
16. Khawaja K, Al-Maini M. Access to pediatric rheumatology care for Juvenile Idiopathic Arthritis in the United Arab Emirates. Pediatr Rheumatol Online J 2017 May 16;15(1):41 [FREE Full text] [doi: 10.1186/s12969-017-0170-4] [Medline: 28511684]
17. Ooi PL, Shek LP. Paediatric rheumatology: a subspecialty in its infancy that is making leaps and bounds. Singapore Med J 2014 May;55(5):242-243 [FREE Full text] [doi: 10.11622/smedj.2014074] [Medline: 24862745]
18. Correll CK, Ditmyer MM, Mehta J, Imundo LF, Klein-Gitelman MS, Monrad SU, et al. 2015 American college of rheumatology workforce study and demand projections of pediatric rheumatology workforce, 2015-2030. Arthritis Care Res (Hoboken) 2020 Oct 27 (forthcoming). [doi: 10.1002/acr.24497] [Medline: 33107674]
19. Berthold E, Månsson B, Kahn R. Outcome in juvenile idiopathic arthritis: a population-based study from Sweden. Arthritis Res Ther 2019 Oct 28;21(1):218 [FREE Full text] [doi: 10.1186/s13075-019-1994-8] [Medline: 31661011]

20. McErlane F, Foster HE, Carrasco R, Baildam EM, Chieng SE, Davidson JE, et al. Trends in paediatric rheumatology referral times and disease activity indices over a ten-year period among children and young people with Juvenile Idiopathic Arthritis: results from the childhood arthritis prospective Study. Rheumatology (Oxford) 2016 Jul;55(7):1225-1234 [FREE Full text] [doi: 10.1093/rheumatology/kew021] [Medline: 27016664]
21. Adriano L, de França Fonteles MM, de Fátima Menezes Azevedo M, Beserra M, Romero N. Medication adherence in patients with juvenile idiopathic arthritis. Rev Bras Reumatol Engl Ed 2017;57(1):23-29 [FREE Full text] [doi: 10.1016/j.rbre.2016.05.004] [Medline: 28137399]
22. Len CA, Miotto e Silva VB, Terreri MT. Importance of adherence in the outcome of juvenile idiopathic arthritis. Curr Rheumatol Rep 2014 Apr;16(4):410. [doi: 10.1007/s11926-014-0410-2] [Medline: 24504596]
23. Sawyer SM, Aroni RA. Self-management in adolescents with chronic illness. What does it mean and how can it be achieved? Med J Aust 2005 Oct 17;183(8):405-409. [doi: 10.5694/j.1326-5377.2005.tb07103.x] [Medline: 16225444]
24. Favier LA, Taylor J, Loiselle Rich K, Jones KB, Vora SS, Harris JG, et al. Barriers to adherence in juvenile idiopathic arthritis: a multicenter collaborative experience and preliminary results. J Rheumatol 2018 May;45(5):690-696 [FREE Full text] [doi: 10.3899/jrheum.171087] [Medline: 29419467]
25. Boland L, Graham ID, Légaré F, Lewis K, Jull J, Shephard A, et al. Barriers and facilitators of pediatric shared decision-making: a systematic review. Implement Sci 2019 Jan 18;14(1):7 [FREE Full text] [doi: 10.1186/s13012-018-0851-5] [Medline: 30658670]
26. Heath G, Farre A, Shaw K. Parenting a child with chronic illness as they transition into adulthood: a systematic review and thematic synthesis of parents' experiences. Patient Educ Couns 2017 Jan;100(1):76-92. [doi: 10.1016/j.pec.2016.08.011] [Medline: 27693084]
27. Coda A, Sculley D, Santos D, Girones X, Brosseau L, Smith D, et al. Harnessing interactive technologies to improve health outcomes in juvenile idiopathic arthritis. Pediatr Rheumatol Online J 2017 May 16;15(1):40 [FREE Full text] [doi: 10.1186/s12969-017-0168-y] [Medline: 28511689]
28. WHO Guideline: Recommendations on Digital Interventions for Health System Strengthening. World Health Organization. 2019. URL: https://www.who.int/reproductivehealth/publications/digital-interventions-health-system-strengthening/en/ [accessed 2022-01-14]
29. Agarwal S, LeFevre AE, Lee J, L'Engle K, Mehl G, Sinha C, WHO mHealth Technical Evidence Review Group. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. BMJ 2016 Mar 17;352:i1174. [doi: 10.1136/bmj.i1174] [Medline: 26988021]
30. Reeder B, David A. Health at hand: a systematic review of smart watch uses for health and wellness. J Biomed Inform 2016 Oct;63:269-276 [FREE Full text] [doi: 10.1016/j.jbi.2016.09.001] [Medline: 27612974]
31. Brigden A, Anderson E, Linney C, Morris R, Parslow R, Serafimova T, et al. Digital behavior change interventions for younger children with chronic health conditions: systematic review. J Med Internet Res 2020 Jul 31;22(7):e16924 [FREE Full text] [doi: 10.2196/16924] [Medline: 32735227]
32. Lau N, Waldbaum S, Parigoris R, O'Daffer A, Walsh C, Colt SF, et al. eHealth and mHealth psychosocial interventions for youths with chronic illnesses: systematic review. JMIR Pediatr Parent 2020 Nov 10;3(2):e22329 [FREE Full text] [doi: 10.2196/22329] [Medline: 33075743]
33. Mosa AS, Yoo I, Sheets L. A systematic review of healthcare applications for smartphones. BMC Med Inform Decis Mak 2012 Jul 10;12:67 [FREE Full text] [doi: 10.1186/1472-6947-12-67] [Medline: 22781312]
34. Griffiths AJ, White CM, Thain PK, Bearne LM. The effect of interactive digital interventions on physical activity in people with inflammatory arthritis: a systematic review. Rheumatol Int 2018 Sep;38(9):1623-1634 [FREE Full text] [doi: 10.1007/s00296-018-4010-8] [Medline: 29556750]
35. Waite-Jones JM, Majeed-Ariss R, Smith J, Stones SR, Van Rooyen V, Swallow V. Young people's, parents', and professionals' views on required components of mobile apps to support self-management of juvenile arthritis: qualitative study. JMIR Mhealth Uhealth 2018 Jan 19;6(1):e25 [FREE Full text] [doi: 10.2196/mhealth.9179] [Medline: 29351898]
36. Johnson KR, Fuchs E, Horvath KJ, Scal P. Distressed and looking for help: internet intervention support for arthritis self-management. J Adolesc Health 2015 Jun;56(6):666-671. [doi: 10.1016/j.jadohealth.2015.02.019] [Medline: 26003583]
37. Park E, Kwon M. Health-related internet use by children and adolescents: systematic review. J Med Internet Res 2018 Apr 03;20(4):e120 [FREE Full text] [doi: 10.2196/jmir.7731] [Medline: 29615385]
38. Villanti AC, Johnson AL, Ilakkuvan V, Jacobs MA, Graham AL, Rath JM. Social media use and access to digital technology in US young adults in 2016. J Med Internet Res 2017 Jun 07;19(6):e196 [FREE Full text] [doi: 10.2196/jmir.7303] [Medline: 28592394]
39. Nikolaou CK, Tay Z, Leu J, Rebello SA, Te Morenga L, Van Dam RM, et al. Young people's attitudes and motivations toward social media and mobile apps for weight control: mixed methods study. JMIR Mhealth Uhealth 2019 Oct 10;7(10):e11205 [FREE Full text] [doi: 10.2196/11205] [Medline: 31603431]
40. Skinner H, Biscope S, Poland B, Goldberg E. How adolescents use technology for health information: implications for health professionals from focus group studies. J Med Internet Res 2003 Dec 18;5(4):e32 [FREE Full text] [doi: 10.2196/jmir.5.4.e32] [Medline: 14713660]

41. Lupton D. Young people's use of digital health technologies in the global north: narrative review. J Med Internet Res 2021 Jan 11;23(1):e18286 [FREE Full text] [doi: 10.2196/18286] [Medline: 33427684]
42. Fay M, Rapley T, Foster H, Pain C, Gerrand C. Can seeding in the clinic reach a wide audience? A proof of concept study on spreading a health message about juvenile idiopathic arthritis using a shareable online video. Interact J Med Res 2016 Feb 22;5(1):e6 [FREE Full text] [doi: 10.2196/ijmr.4608] [Medline: 26903485]
43. Armbrust W, Bos JJ, Cappon J, van Rossum MA, Sauer PJ, Wulffraat N, et al. Design and acceptance of Rheumates@Work, a combined internet-based and in person instruction model, an interactive, educational, and cognitive behavioral program for children with juvenile idiopathic arthritis. Pediatr Rheumatol Online J 2015 Jul 23;13:31 [FREE Full text] [doi: 10.1186/s12969-015-0029-5] [Medline: 26202161]
44. Cai RA, Beste D, Chaplin H, Varakliotis S, Suffield L, Josephs F, et al. Developing and evaluating JIApp: acceptability and usability of a smartphone app system to improve self-management in young people with juvenile idiopathic arthritis. JMIR Mhealth Uhealth 2017 Aug 15;5(8):e121 [FREE Full text] [doi: 10.2196/mhealth.7229] [Medline: 28811270]
45. Lee RR, Rashid A, Ghio D, Thomson W, Cordingley L. "Seeing Pain Differently": a qualitative investigation into the differences and similarities of pain and rheumatology specialists' interpretation of multidimensional mobile health pain data from children and young people with juvenile idiopathic arthritis. JMIR Mhealth Uhealth 2019 Jul 02;7(7):e12952 [FREE Full text] [doi: 10.2196/12952] [Medline: 31267979]
46. Butler S, Sculley D, Santos DS, Fellas A, Gironès X, Singh-Grewal D, et al. Usability of eHealth and mobile health interventions by young people living with juvenile idiopathic arthritis: systematic review. JMIR Pediatr Parent 2020 Dec 01;3(2):e15833 [FREE Full text] [doi: 10.2196/15833] [Medline: 33258786]
47. Leigh S, Ashall-Payne L, Andrews T. Barriers and facilitators to the adoption of mobile health among health care professionals from the united kingdom: discrete choice experiment. JMIR Mhealth Uhealth 2020 Jul 06;8(7):e17704 [FREE Full text] [doi: 10.2196/17704] [Medline: 32628118]
48. Monitoring and Evaluating Digital Health Interventions: A Practical Guide to Conducting Research and Assessment. World Health Organization. 2016. URL: https://www.who.int/reproductivehealth/publications/mhealth/digital-health-interventions/en/ [accessed 2022-01-14]
49. Health of children: profile of children. Australian Institute of Health and Welfare. 2020. URL: https://www.aihw.gov.au/reports/australias-health/health-of-children [accessed 2022-01-14]
50. Sawyer S, McNeil R, Francis K, Matskarofski J, Patton G, Bhutta Z, et al. The age of paediatrics. The Lancet Child & Adolescent Health 2019 Nov;3(11):822-830 [FREE Full text] [doi: 10.1016/s2352-4642(19)30266-4]
51. Infants (0-1 year of age). Centers for Disease Control and Prevention. URL: https://www.cdc.gov/ncbddd/childdevelopment/positiveparenting/infants.html [accessed 2021-11-30]
52. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
53. Butler S, Sculley D, Santos DS, Fellas A, Gironès X, Singh-Grewal D, et al. Usability of eHealth and mobile health interventions by young people living with juvenile idiopathic arthritis: systematic review. JMIR Pediatr Parent 2020 Dec 01;3(2):e15833 [FREE Full text] [doi: 10.2196/15833] [Medline: 33258786]
54. Mesh. National Centre for Biotechnology Information. URL: https://www.ncbi.nlm.nih.gov/mesh/ [accessed 2022-01-14]
55. Covidence. Cochrane Community. URL: https://community.cochrane.org/help/tools-and-software/covidence [accessed 2022-01-09]
56. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health 1998 Jun;52(6):377-384 [FREE Full text] [doi: 10.1136/jech.52.6.377] [Medline: 9764259]
57. Checklist for measuring study quality. National Collaborating Centre for Methods and Tools. URL: https://www.nccmt.ca/knowledge-repositories/search/9 [accessed 2022-01-14]
58. Ryan R. Cochrane Consumers and Communication Review Group: data synthesis and analysis. Cochrane Consumers and Communication Review Group. URL: https://cccrg.cochrane.org/sites/cccrg.cochrane.org/files/public/uploads/Analysis.pdf [accessed 2022-01-14]
59. RevMan for non-Cochrane reviews. Cochrane Training. URL: https://training.cochrane.org/online-learning/core-software-cochrane-reviews/revman/revman-non-cochrane-reviews [accessed 2022-01-14]
60. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. Res Synth Methods 2010 Apr;1(2):97-111. [doi: 10.1002/jrsm.12] [Medline: 26061376]
61. Tufanaru C, Munn Z, Stephenson M, Aromataris E. Fixed or random effects meta-analysis? Common methodological issues in systematic reviews of effectiveness. Int J Evid Based Healthc 2015 Sep;13(3):196-207. [doi: 10.1097/XEB.0000000000000065] [Medline: 26355603]
62. Brydges CR. Effect size guidelines, sample size calculations, and statistical power in gerontology. Innov Aging 2019 Aug;3(4):igz036 [FREE Full text] [doi: 10.1093/geroni/igz036] [Medline: 31528719]
63. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003 Sep 06;327(7414):557-560 [FREE Full text] [doi: 10.1136/bmj.327.7414.557] [Medline: 12958120]

64. Schünemann HJ, Vist GE, Higgins JPT, Santesso N, Deeks JJ, Glasziou P, et al. Chapter 15: interpreting results, drawing conclusions. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, editors. Cochrane Handbook for Systematic Reviews of Interventions version 6.2. London, England: Cochrane; 2021.
65. Armbrust W, Bos GJ, Wulffraat NM, van Brussel M, Cappon J, Dijkstra PU, et al. Internet program for physical activity and exercise capacity in children with juvenile idiopathic arthritis: a multicenter randomized controlled trial. Arthritis Care Res (Hoboken) 2017 Jul;69(7):1040-1049. [doi: 10.1002/acr.23100] [Medline: 27696793]
66. Connelly M, Schanberg LE, Ardoin S, Blakley M, Carrasco R, Chira P, et al. Multisite randomized clinical trial evaluating an online self-management program for adolescents with juvenile idiopathic arthritis. J Pediatr Psychol 2019 Apr 01;44(3):363-374 [FREE Full text] [doi: 10.1093/jpepsy/jsy066] [Medline: 30204919]
67. Doeleman M, de Roock S, Buijsse N, Klein M, Bonsel GJ, Seyfert-Margolis V, et al. Monitoring patients with juvenile idiopathic arthritis using health-related quality of life. Pediatr Rheumatol Online J 2021 Mar 22;19(1):40 [FREE Full text] [doi: 10.1186/s12969-021-00527-z] [Medline: 33752695]
68. Haverman L, van Rossum MA, van Veenendaal M, van den Berg JM, Dolman KM, Swart J, et al. Effectiveness of a web-based application to monitor health-related quality of life. Pediatrics 2013 Feb;131(2):e533-e543. [doi: 10.1542/peds.2012-0958] [Medline: 23296436]
69. Heale LD, Dover S, Goh YI, Maksymiuk VA, Wells GD, Feldman BM. A wearable activity tracker intervention for promoting physical activity in adolescents with juvenile idiopathic arthritis: a pilot study. Pediatr Rheumatol Online J 2018 Oct 22;16(1):66 [FREE Full text] [doi: 10.1186/s12969-018-0282-5] [Medline: 30348203]
70. Lalloo C, Harris LR, Hundert AS, Berard R, Cafazzo J, Connelly M, et al. The iCanCope pain self-management application for adolescents with juvenile idiopathic arthritis: a pilot randomized controlled trial. Rheumatology (Oxford) 2021 Jan 05;60(1):196-206. [doi: 10.1093/rheumatology/keaa178] [Medline: 32613229]
71. Lee RR, Shoop-Worrall S, Rashid A, Thomson W, Cordingley L. "Asking too much?": randomized n-of-1 trial exploring patient preferences and measurement reactivity to frequent use of remote multidimensional pain assessments in children and young people with juvenile idiopathic arthritis. J Med Internet Res 2020 Jan 30;22(1):e14503 [FREE Full text] [doi: 10.2196/14503] [Medline: 32012051]
72. Lelieveld OT, Armbrust W, Geertzen JH, de Graaf I, van Leeuwen MA, Sauer PJ, et al. Promoting physical activity in children with juvenile idiopathic arthritis through an internet-based program: results of a pilot randomized controlled trial. Arthritis Care Res (Hoboken) 2010 May;62(5):697-703 [FREE Full text] [doi: 10.1002/acr.20085] [Medline: 20191468]
73. Stinson JN, Lalloo C, Hundert AS, Campillo S, Cellucci T, Dancey P, et al. Teens taking charge: a randomized controlled trial of a web-based self-management program with telephone support for adolescents with juvenile idiopathic arthritis. J Med Internet Res 2020 Jul 29;22(7):e16234 [FREE Full text] [doi: 10.2196/16234] [Medline: 32723728]
74. Stinson J, Ahola Kohut S, Forgeron P, Amaria K, Bell M, Kaufman M, et al. The iPeer2Peer Program: a pilot randomized controlled trial in adolescents with Juvenile Idiopathic Arthritis. Pediatr Rheumatol Online J 2016 Sep 02;14(1):48 [FREE Full text] [doi: 10.1186/s12969-016-0108-2] [Medline: 27590668]
75. Stinson JN, Jibb LA, Lalloo C, Feldman BM, McGrath PJ, Petroz GC, et al. Comparison of average weekly pain using recalled paper and momentary assessment electronic diary reports in children with arthritis. Clin J Pain 2014 Dec;30(12):1044-1050. [doi: 10.1097/AJP.0000000000000072] [Medline: 24535055]
76. Stinson JN, Connelly M, Jibb LA, Schanberg LE, Walco G, Spiegel LR, et al. Developing a standardized approach to the assessment of pain in children and youth presenting to pediatric rheumatology providers: a Delphi survey and consensus conference process followed by feasibility testing. Pediatr Rheumatol Online J 2012 Apr 10;10(1):7 [FREE Full text] [doi: 10.1186/1546-0096-10-7] [Medline: 22490427]
77. Stinson JN, McGrath PJ, Hodnett ED, Feldman BM, Duffy CM, Huber AM, et al. An internet-based self-management program with telephone support for adolescents with arthritis: a pilot randomized controlled trial. J Rheumatol 2010 Sep;37(9):1944-1952. [doi: 10.3899/jrheum.091327] [Medline: 20595280]
78. Stinson JN, Petroz GC, Stevens BJ, Feldman BM, Streiner D, McGrath PJ, et al. Working out the kinks: testing the feasibility of an electronic pain diary for adolescents with arthritis. Pain Res Manag 2008;13(5):375-382 [FREE Full text] [doi: 10.1155/2008/326389] [Medline: 18958308]
79. Stinson JN, Stevens BJ, Feldman BM, Streiner D, McGrath PJ, Dupuis A, et al. Construct validity of a multidimensional electronic pain diary for adolescents with arthritis. Pain 2008 Jun;136(3):281-292. [doi: 10.1016/j.pain.2007.07.002] [Medline: 17723279]
80. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J, International League of Associations for Rheumatology. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol 2004 Feb;31(2):390-392. [Medline: 14760812]
81. Classification of digital health interventions v1.0: a shared language to describe the uses of digital technology for health. World Health Organization. 2018. URL: https://apps.who.int/iris/handle/10665/260480 [accessed 2022-01-14]
82. Rashid A, Cordingley L, Carrasco R, Foster HE, Baildam EM, Chieng A, et al. Patterns of pain over time among children with juvenile idiopathic arthritis. Arch Dis Child 2018 May;103(5):437-443 [FREE Full text] [doi: 10.1136/archdischild-2017-313337] [Medline: 29175824]

83. Walker SM. Overview of neurodevelopment and pain research, possible treatment targets. Best Pract Res Clin Rheumatol 2014 Apr;28(2):213-228. [doi: 10.1016/j.berh.2014.03.007] [Medline: 24974059]
84. Breivik H, Borchgrevink P, Allen S, Rosseland L, Romundstad L, Hals EK, et al. Assessment of pain. Br J Anaesth 2008 Jul;101(1):17-24 [FREE Full text] [doi: 10.1093/bja/aen103] [Medline: 18487245]
85. Logan DE, Conroy C, Sieberg CB, Simons LE. Changes in willingness to self-manage pain among children and adolescents and their parents enrolled in an intensive interdisciplinary pediatric pain treatment program. Pain 2012 Sep;153(9):1863-1870 [FREE Full text] [doi: 10.1016/j.pain.2012.05.027] [Medline: 22749194]
86. Stone AA, Broderick JE, Kaell AT. Single momentary assessments are not reliable outcomes for clinical trials. Contemp Clin Trials 2010 Sep;31(5):466-472 [FREE Full text] [doi: 10.1016/j.cct.2010.05.006] [Medline: 20580945]
87. Heron E, Everhart R, McHale S, Smyth JM. Using mobile-technology-based Ecological Momentary Assessment (EMA) methods with youth: a systematic review and recommendations. J Pediatr Psychol 2017 Nov 01;42(10):1087-1107. [doi: 10.1093/jpepsy/jsx078] [Medline: 28475765]
88. Stone AA, Broderick JE, Shiffman SS, Schwartz JE. Understanding recall of weekly pain from a momentary assessment perspective: absolute agreement, between- and within-person consistency, and judged change in weekly pain. Pain 2004 Jan;107(1-2):61-69. [doi: 10.1016/j.pain.2003.09.020] [Medline: 14715390]
89. McBain H, Shipley M, Newman S. The impact of self-monitoring in chronic illness on healthcare utilisation: a systematic review of reviews. BMC Health Serv Res 2015 Dec 18;15:565 [FREE Full text] [doi: 10.1186/s12913-015-1221-5] [Medline: 26684011]
90. Walker RC, Tong A, Howard K, Palmer SC. Patient expectations and experiences of remote monitoring for chronic diseases: systematic review and thematic synthesis of qualitative studies. Int J Med Inform 2019 Apr;124:78-85 [FREE Full text] [doi: 10.1016/j.ijmedinf.2019.01.013] [Medline: 30784430]
91. Hvitfeldt H, Carli C, Nelson EC, Mortenson DM, Ruppert BA, Lindblad S. Feed forward systems for patient participation and provider support: adoption results from the original US context to Sweden and beyond. Qual Manag Health Care 2009;18(4):247-256. [doi: 10.1097/QMH.0b013e3181bee32e] [Medline: 19851232]
92. Grande S, Longacre M, Palmblad K, Montan M, Berquist R, Hager A, et al. Empowering young people living with juvenile idiopathic arthritis to better communicate with families and care teams: content analysis of semistructured interviews. JMIR Mhealth Uhealth 2019 Feb 22;7(2):e10401 [FREE Full text] [doi: 10.2196/10401] [Medline: 30794202]
93. Ammerlaan JJ, Scholtus LW, Drossaert CH, van Os-Medendorp H, Prakken B, Kruize AA, et al. Feasibility of a website and a hospital-based online portal for young adults with juvenile idiopathic arthritis. JMIR Res Protoc 2015 Aug 14;4(3):e102 [FREE Full text] [doi: 10.2196/resprot.3952] [Medline: 26276373]
94. Geerards D, Pusic A, Hoogbergen M, van der Hulst R, Sidey-Gibbons C. Computerized quality of life assessment: a randomized experiment to determine the impact of individualized feedback on assessment experience. J Med Internet Res 2019 Jul 11;21(7):e12212 [FREE Full text] [doi: 10.2196/12212] [Medline: 31298217]
95. Beacham BL, Deatrick JA. Children with chronic conditions: perspectives on condition management. J Pediatr Nurs 2015;30(1):25-35 [FREE Full text] [doi: 10.1016/j.pedn.2014.10.011] [Medline: 25458105]
96. Bos GJ, Lelieveld OT, Armbrust W, Sauer PJ, Geertzen JH, Dijkstra PU. Physical activity in children with Juvenile Idiopathic Arthritis compared to controls. Pediatr Rheumatol Online J 2016 Jul 07;14(1):42 [FREE Full text] [doi: 10.1186/s12969-016-0102-8] [Medline: 27387754]
97. Maggio AB, Hofer MF, Martin XE, Marchand LM, Beghetti M, Farpour-Lambert NJ. Reduced physical activity level and cardiorespiratory fitness in children with chronic diseases. Eur J Pediatr 2010 Oct;169(10):1187-1193. [doi: 10.1007/s00431-010-1199-2] [Medline: 20411275]
98. Kuntze G, Nesbitt C, Whittaker JL, Nettel-Aguirre A, Toomey C, Esau S, et al. Exercise therapy in juvenile idiopathic arthritis: a systematic review and meta-analysis. Arch Phys Med Rehabil 2018 Jan;99(1):178-93.e1. [doi: 10.1016/j.apmr.2017.05.030] [Medline: 28729171]
99. Bouaddi I, Rostom S, El Badri D, Hassani A, Chkirate B, Amine B, et al. Impact of juvenile idiopathic arthritis on schooling. BMC Pediatr 2013 Jan 07;13:2 [FREE Full text] [doi: 10.1186/1471-2431-13-2] [Medline: 23289498]
100. Laila K, Haque M, Islam MM, Islam MI, Talukder MK, Rahman SA. Impact of juvenile idiopathic arthritis on school attendance and performance. Am J Clin Exper Med 2016 Nov;4(6):185-190. [doi: 10.11648/j.ajcem.20160406.15]
101. Feinstein AB, Cohen LL, Masuda A, Griffin AT, Gamwell KL, Stiles MT, et al. Yoga intervention for an adolescent with juvenile idiopathic arthritis: a case study. Adv Mind Body Med 2018;32(1):13-20. [Medline: 29406303]
102. Sandstedt E, Fasth A, Eek MN, Beckung E. Muscle strength, physical fitness and well-being in children and adolescents with juvenile idiopathic arthritis and the effect of an exercise programme: a randomized controlled trial. Pediatr Rheumatol Online J 2013 Feb 22;11(1):7 [FREE Full text] [doi: 10.1186/1546-0096-11-7] [Medline: 23432796]
103. Feldman D, de Civita M, Dobkin P, Malleson P, Meshefedjian G, Duffy C. Perceived adherence to prescribed treatment in juvenile idiopathic arthritis over a one-year period. Arthritis Rheum 2007 Mar 15;57(2):226-233 [FREE Full text] [doi: 10.1002/art.22534] [Medline: 17330298]
104. Heapy AA, Higgins DM, Cervone D, Wandner L, Fenton BT, Kerns RD. A systematic review of technology-assisted self-management interventions for chronic pain: looking across treatment modalities. Clin J Pain 2015 Jun;31(6):470-492. [doi: 10.1097/AJP.0000000000000185] [Medline: 25411862]

105. Chapman L, Brooks C, Lawson J, Russell C, Adams J. Accessibility of online self-management support websites for people with osteoarthritis: a text content analysis. Chronic Illn 2019 Mar;15(1):27-40. [doi: 10.1177/1742395317746471] [Medline: 29254372]
106. Devine T, Broderick J, Harris LM, Wu H, Hilfiker SW. Making quality health websites a national public health priority: toward quality standards. J Med Internet Res 2016 Aug 02;18(8):e211 [FREE Full text] [doi: 10.2196/jmir.5999] [Medline: 27485512]
107. Ledel Solem IK, Varsi C, Eide H, Kristjansdottir OB, Mirkovic J, Børøsund E, et al. Patients' needs and requirements for eHealth pain management interventions: qualitative study. J Med Internet Res 2019 Apr 01;21(4):e13205 [FREE Full text] [doi: 10.2196/13205] [Medline: 30877780]
108. Henrickson M. Policy challenges for the pediatric rheumatology workforce: Part III. the international situation. Pediatr Rheumatol Online J 2011 Sep 12;9:26 [FREE Full text] [doi: 10.1186/1546-0096-9-26] [Medline: 21910871]
109. Henrickson M. Policy challenges for the pediatric rheumatology workforce: Part II. Health care system delivery and workforce supply. Pediatr Rheumatol Online J 2011;9:24 [FREE Full text] [doi: 10.1186/1546-0096-9-23] [Medline: 21843335]
110. Pedersen DH, Mansourvar M, Sortsø C, Schmidt T. Predicting dropouts from an electronic health platform for lifestyle interventions: analysis of methods and predictors. J Med Internet Res 2019 Sep 04;21(9):e13617 [FREE Full text] [doi: 10.2196/13617] [Medline: 31486409]
111. Winkelman TN, Caldwell MT, Bertram B, Davis MM. Promoting health literacy for children and adolescents. Pediatrics 2016 Dec;138(6):e20161937. [doi: 10.1542/peds.2016-1937] [Medline: 27940702]
112. Bröder J, Okan O, Bauer U, Bruland D, Schlupp S, Bollweg TM, et al. Health literacy in childhood and youth: a systematic review of definitions and models. BMC Public Health 2017 Apr 26;17(1):361 [FREE Full text] [doi: 10.1186/s12889-017-4267-y] [Medline: 28441934]
113. Miller TA. Health literacy and adherence to medical treatment in chronic and acute illness: a meta-analysis. Patient Educ Couns 2016 Jul;99(7):1079-1086 [FREE Full text] [doi: 10.1016/j.pec.2016.01.020] [Medline: 26899632]
114. Bondaronek P, Alkhaldi G, Slee A, Hamilton FL, Murray E. Quality of publicly available physical activity apps: review and content analysis. JMIR Mhealth Uhealth 2018 Mar 21;6(3):e53 [FREE Full text] [doi: 10.2196/mhealth.9069] [Medline: 29563080]
115. Nicolson PJ, Hinman RS, Wrigley TV, Stratford PW, Bennell KL. Self-reported home exercise adherence: a validity and reliability study using concealed accelerometers. J Orthop Sports Phys Ther 2018 Dec;48(12):943-950. [doi: 10.2519/jospt.2018.8275] [Medline: 30053792]
116. Loprinzi P, Cardinal B. Measuring children's physical activity and sedentary behaviors. J Exerc Sci Fitness 2011 Jun;9(1):15-23 [FREE Full text] [doi: 10.1016/s1728-869x(11)60002-6]
117. Armbrust W, Bos GJ, Geertzen JH, Sauer PJ, Dijkstra PU, Lelieveld OT. Measuring physical activity in juvenile idiopathic arthritis: activity diary versus accelerometer. J Rheumatol 2017 Aug;44(8):1249-1256. [doi: 10.3899/jrheum.160671] [Medline: 28365580]
118. Mackintosh KA, Chappel SE, Salmon J, Timperio A, Ball K, Brown H, et al. Parental perspectives of a wearable activity tracker for children younger than 13 years: acceptability and usability study. JMIR Mhealth Uhealth 2019 Nov 04;7(11):e13858 [FREE Full text] [doi: 10.2196/13858] [Medline: 31682585]
119. Borthwick AC, Anderson CL, Finsness ES, Foulger TS. Special article personal wearable technologies in education: value or villain? J Digit Learn Teacher Educ 2015 Jul 02;31(3):85-92. [doi: 10.1080/21532974.2015.1021982]

Abbreviations

CG: control group
EQ-5D-Y-5L: EuroQol 5-dimensional youth 5-level
HRQoL: health-related quality of life
IG: intervention group
JIA: juvenile idiopathic arthritis
mHealth: mobile health
MVPA: moderate to vigorous physical activity
PAL: physical activity level
PedsQL: Pediatric Quality of Life Inventory
PR: pediatric rheumatologist
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
SMD: standardized mean difference

Edited by R Kukafka, G Eysenbach; submitted 22.05.21; peer-reviewed by R Bajpai, E Sadeghi-Demneh, F Ghezelbash, S Shoop-Worrall, J McDonagh; comments to author 30.06.21; revised version received 10.09.21; accepted 08.11.21; published 02.02.22.

Please cite as:
Butler S, Sculley D, Santos D, Fellas A, Gironès X, Singh-Grewal D, Coda A
Effectiveness of eHealth and mHealth Interventions Supporting Children and Young People Living With Juvenile Idiopathic Arthritis: Systematic Review and Meta-analysis
J Med Internet Res 2022;24(2):e30457
URL: https://www.jmir.org/2022/2/e30457
doi:10.2196/30457
PMID:35107431

Review

Digital Approaches to Automated and Machine Learning Assessments of Hearing: Scoping Review

Jan-Willem Wasmann[1*], MSc; Leontien Pragt[1*], MSc; Robert Eikelboom[2,3,4], PhD; De Wet Swanepoel[2,3,4], PhD

[1]Department of Otorhinolaryngology, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Centre, Nijmegen, Netherlands

[2]Ear Science Institute Australia, Subiaco, Australia

[3]Ear Sciences Centre, Medical School, The University of Western Australia, Perth, Australia

[4]Department of Speech-Language Pathology and Audiology, University of Pretoria, Pretoria, South Africa

[*]these authors contributed equally

Corresponding Author:
Jan-Willem Wasmann, MSc
Department of Otorhinolaryngology, Donders Institute for Brain, Cognition and Behaviour
Radboud University Medical Centre
Philips van Leydenlaan 15
Nijmegen, 6500 HB
Netherlands
Phone: 31 024 361 04 2
Email: Jan-Willem.Wasmann@radboudumc.nl

Abstract

Background: Hearing loss affects 1 in 5 people worldwide and is estimated to affect 1 in 4 by 2050. Treatment relies on the accurate diagnosis of hearing loss; however, this first step is out of reach for >80% of those affected. Increasingly automated approaches are being developed for self-administered digital hearing assessments without the direct involvement of professionals.

Objective: This study aims to provide an overview of digital approaches in automated and machine learning assessments of hearing using pure-tone audiometry and to focus on the aspects related to accuracy, reliability, and time efficiency. This review is an extension of a 2013 systematic review.

Methods: A search across the electronic databases of PubMed, IEEE, and Web of Science was conducted to identify relevant reports from the peer-reviewed literature. Key information about each report's scope and details was collected to assess the commonalities among the approaches.

Results: A total of 56 reports from 2012 to June 2021 were included. From this selection, 27 unique automated approaches were identified. Machine learning approaches require fewer trials than conventional threshold-seeking approaches, and personal digital devices make assessments more affordable and accessible. Validity can be enhanced using digital technologies for quality surveillance, including noise monitoring and detecting inconclusive results.

Conclusions: In the past 10 years, an increasing number of automated approaches have reported similar accuracy, reliability, and time efficiency as manual hearing assessments. New developments, including machine learning approaches, offer features, versatility, and cost-effectiveness beyond manual audiometry. Used within identified limitations, automated assessments using digital devices can support task-shifting, self-care, telehealth, and clinical care pathways.

(J Med Internet Res 2022;24(2):e32581) doi:10.2196/32581

KEYWORDS

audiology; automated audiometry; automatic audiometry; automation; digital health technologies; digital hearing health care; machine learning; remote care; self-administered audiometry; self-assessment audiometry; user-operated audiometry; digital health; hearing loss; digital hearing; digital devices; mobile phone; telehealth

Introduction

Background

Hearing loss affects 1.5 billion persons worldwide and is expected to increase by another billion by 2050 [1,2]. Hearing testing is the first step toward appropriate and timely treatment. Unfortunately, most persons affected with hearing loss are unable to access hearing assessments, with less than one hearing health professional for every million people in regions such as Africa [2,3]. Increasingly automated approaches (all aspects of the method associated with automated audiometry), including machine learning, are being developed and made available to provide self-administered hearing assessments. The term *automated audiometry* refers to all hearing tests that are self-administered from the point the test starts. More specifically, in this review, we define automated audiometry as calibrated pure-tone threshold audiometry in any setting (ie, hearing health care, occupational health, and community settings) that is self-administered from the point the test starts. Machine learning refers to model-based approaches that learn from examples (data) instead of being programmed with rules [4]. As the direct involvement of professionals is not required, automated approaches enable health care pathways with the potential to increase accessibility, efficiency, and scalability. Digital (health) technologies, including apps, smartphones, tablets, and wearables, can acquire data remotely; expand the reach and precision of clinicians; and facilitate more personalized hearing health care within a network of distributed expertise [5,6]. Recent examples of automated hearing assessments include clinical grade and consumer-grade applications [7]. General global health trends suggest that increased availability of diagnostic tools could lower health care costs and improve quality of life [8]. For example, in Parkinson disease, remote care based on wearables provides ecologically valid methods for monitoring and evaluating symptoms [9,10]. In tuberculosis screening in low-resource settings, an automated diagnosis can increase the sensitivity of identifying persons at risk while reducing costs [11]. Self-assessment using eHealth vision tools improves access to diagnosis and facilitates timely diagnosis, although consistent criteria for referring to the clinical pathway and validity and reliability of eHealth tools are still a concern [12].

Timely detection and treatment of hearing loss are essential to enable optimal outcomes and quality of life across the life span [2]. Untreated hearing loss restricts language development and educational potential in children and is associated with a more rapid cognitive decline in adults [13]. It may lead to social isolation, lower socioeconomic status, increased social disparities, and decreased health, resulting in lower quality of life at the individual level and substantial costs at the community level [14,15]. Importantly, treating hearing loss in midlife has been identified as the largest potentially modifiable risk factor for developing dementia in later life [16]. The global annual cost of untreated hearing loss is US $980 million [14]. Global health investment models indicate a significant return on investment in both hearing diagnosis and treatment [2]. The capacity of the entire clinical pathway should be increased as a bottleneck looms if the accessibility of diagnosis is increased faster than the availability of affordable treatment and rehabilitation.

Automated self-test options are important for detecting and diagnosing hearing loss to direct timely and appropriate treatments. The overwhelming majority of treatments are for permanent age-related and noise-induced hearing loss; however, a significant portion of the population requires medical treatment for hearing loss [1]. The onset of the COVID-19 pandemic has further emphasized the importance of self-testing approaches [17,18]. Automation on digital devices is a powerful enabler of alternative diagnostic pathways that can include home-based testing, low-touch service models outside traditional clinic settings, and decentralized community-based models that rely on task shifting to minimally trained facilitators [19].

Automation in hearing assessment is not a new concept and dates back to >7 decades [20]. In recent years, it has resurged with the convergence of digital technologies and machine learning approaches. The primary tool for hearing assessment is pure-tone audiometry, which describes the degree of hearing loss relative to normal hearing, expressed in decibels hearing level across specific frequencies (125-8000 Hz). Pure-tone audiometry can also differentiate the type of hearing loss, that is, sensorineural or conductive, when bone conduction and air conduction transducers are used. Machine learning–based threshold-seeking approaches, known as Bayesian active learning, have demonstrated their potential to optimize efficiency and increase the precision of automated hearing assessments [21]. The increased efficiency comes from the ability of these methods to target trials to those areas of the frequency space where the estimation has the greatest uncertainty [22,23].

Objective

In 2013, a systematic review that included 29 reports on automated audiometry showed that automated procedures have comparable accuracy with that of manual procedures when performing air conduction audiometry. Although a few validated automated procedures that included automated bone conduction audiometry had been reported, machine learning–based audiometry approaches had not been reported yet, and approaches were rarely validated in children or hard-to-test populations [24]. Since 2013, there has been significant work and innovation in this area, which calls for an update and extension of the previous review. This study aims to provide the current status of automation and machine learning approaches in hearing assessment using validated pure-tone audiometry with potential indicators of accuracy, reliability, and efficiency of these approaches.

Methods

We conducted a systematic scoping review of the peer-reviewed literature on automated and machine learning approaches to validate pure-tone threshold audiometry using digital technologies by considering accuracy, reliability, and efficiency. This review followed the methodological framework outlined by Arksey and O'Malley [25].

Identifying Potentially Relevant Records

A search across the electronic databases of PubMed, IEEE, and Web of Science was conducted to identify relevant reports from the peer-reviewed literature. Complementary and redundant search terms were applied to ensure thorough coverage and cross-checking of the search findings. In the PubMed database, medical subject headings and relevant keywords were collected to determine all records related to the study aim. The following synonyms of, and closely related terms to, automated audiometry were used: automatic audiometry, self-administered audiometry, self-assessment audiometry, and user-operated audiometry. The complete set of terms and the applied search strategy are provided in Multimedia Appendix 1. The IEEE database is engineering oriented, and only relevant keywords based on audiometry were used, as it was assumed that any result in audiometry would be highly associated with automated audiometry. The Web of Science database is known to index the PubMed and IEEE databases and was explored using search terms similar to the PubMed search. After preliminary explorations to identify appropriate keywords, we conducted a search on July 8, 2020, and updated it on January 12, 2021, and July 6, 2021. The search included all reports that met the inclusion criteria published from January 1, 2012, to June 30, 2021. The start date was chosen as we regard this scoping review as an extension and generalization of a previous (systematic) review by Mahomed et al [24], which included studies up to July 20, 2012.

Selecting Relevant Records

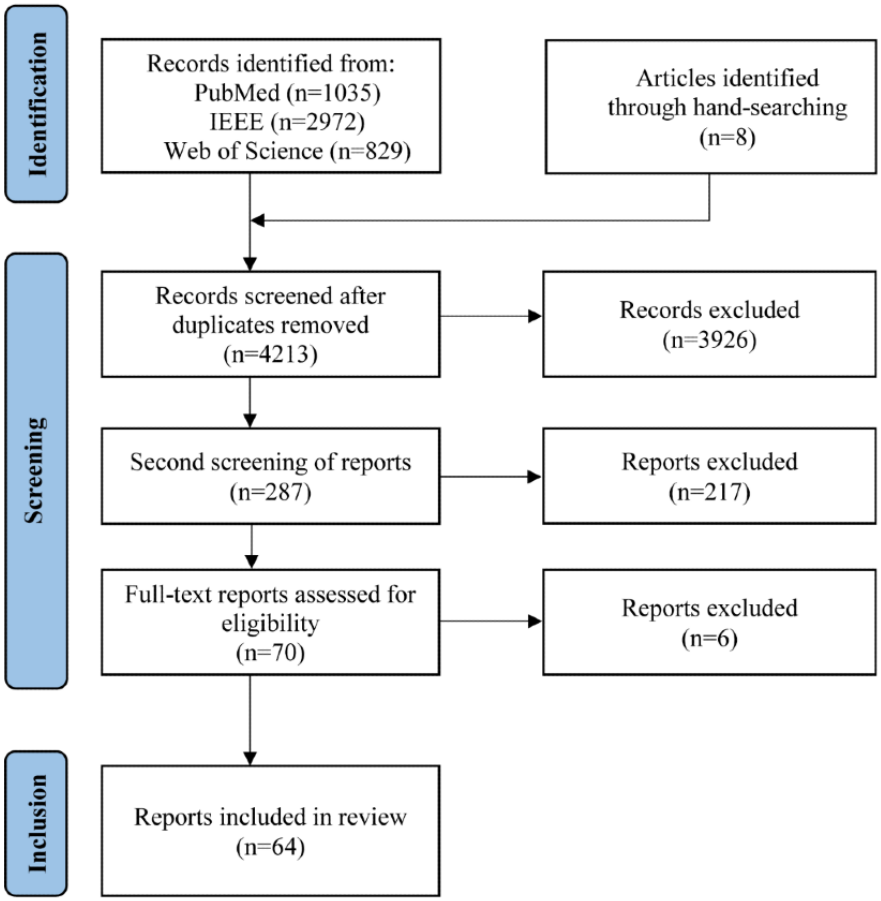
Reports had to meet the following three inclusion criteria: (1) the report had to be about automated or machine learning and pure-tone frequency-specific threshold audiometry, (2) it had to be written in English, and (3) the automated threshold audiometry had to be compared against the gold standard or reasonable standard. The gold standard is defined as manual audiometry in a sound booth according to the International Organization for Standardization standards. Automated audiometry also needed to be performed inside a sound booth, and the results needed to be compared with the gold standard. A reasonable standard for validation was defined as either a within-subject comparison between the gold standard and the automated audiometry in an unconventional setting (eg, a quiet room) or a within-subject comparison between a validated automated audiometry approach and an experimental approach of audiometry in the same unconventional setting.

We excluded reports on screening audiometry (eg, provided pass or refer as an outcome) rather than threshold audiometry, review papers, and studies reporting approaches that were not compared with the gold or reasonable reference standard.

The first phase of screening was based on the title. If the title indicated that content was within the scope of the research question (ie, automated or machine learning approaches in diagnostic hearing assessment), the report was included in the second screening phase. In the second phase, the abstracts of the remaining reports were assessed using the inclusion and exclusion criteria stated earlier.

Two researchers (LP and JWW) conducted the abstract screening. They were blinded from each other to avoid confirmation bias. After the screening, the researchers discussed any disagreements to reach an agreement. When in doubt, the report was admitted to the third, full-text review phase. In this phase, all the remaining reports were reviewed in full to determine whether the inclusion criteria were met. As can be seen in the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram (Figure 1), the resulting selection of reports was complemented by additional reports. After some reports were clustered as having identical approaches (explained in *Collating Approaches, Summarizing, and Reporting the Results*), additional reports were added to avoid missing validation data of these clustered approaches. These additional reports were published before the inclusion date criteria (from before January 1, 2012) or did not appear in the search and were added based on the reference lists of the already included reports.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of the screening process.



Extracting Data Items

A template for grading the reports was agreed upon by all the authors (Multimedia Appendix 2 [26]). Two researchers (LP and JWW) independently extracted information directly relevant to the scoping review question. In cases of disagreement, a consensus was reached after discussion between the 2 researchers. The compulsory data fields were test frequency and intensity range; response method; test equipment, including the type of transducers; calibration; hardware; test quality control; accuracy; reliability; efficiency; validation; and test population. In the report by Mahomed et al [24], the accuracy and reliability of manual and automated approaches demonstrated equivalent performances. Time efficiency had primarily been reported by comparing the testing times of manual and automated audiometry [27-29]. The reports on machine learning audiometry explicitly used the number of trials or stimuli needed to converge to a certain precision (eg, 5 dB) as a performance outcome [23,29]. Therefore, we added time efficiency as a necessary parameter. Where available, accuracy and reliability were expressed in decibels using the overall root mean square deviation (RMSD) between the automated approach and the gold (or reasonable) standard. On the basis of the study by Margolis et al [30] and the minimum acceptable accuracy recommended by clinical guidelines [31], RMSD values of 6 dB and 10 dB were chosen as criteria for desired and minimal accuracy, respectively. To establish a benchmark for an acceptable test duration, the mean testing time for conventional manual bilateral audiometry (air 7 and bone 5 frequencies) was estimated (Multimedia Appendix 3 [27-29,31,34,38]). For manual bilateral air conduction, based on the benchmark measurement times, a mean testing time of 5 to 10 minutes was considered acceptable, and for manual bilateral air and bone conduction, 10 to 20 minutes was considered acceptable. If testing times exceeded these ranges by >5 minutes, the time efficiency was assessed as a potential issue.

Data collected from the reports provided key information about the scope and details of each report, enabling the authors to assess commonalities between the approaches.

Collating Approaches, Summarizing, and Reporting the Results

When multiple reports described the same underlying approach, these reports were pooled into one approach cluster. The first report describing an approach and subsequent studies that validated or extended the approach were included. The name of the approach, citations to the initial report, or common authorships were used to cluster the reports. The grading table was completed for each cluster separately to provide a structure for the subsequent content analysis. In the last part of the grading table, under the heading *Validation Approach*, all validation studies are described together. For every approach cluster, a key contribution to the audiological field was derived from the associated reports. A key contribution is a finding or claim made by the authors significant to the approach in general, stated in

either the conclusion or the discussion section of a report in accordance with their objective.

Results

Overview

A total of 64 reports were included in this study. Of the 64 reports, 56 (88%) were included according to the inclusion and exclusion criteria, and 8 (13%) were added to the approach clusters. After clustering identical approaches, 27 approach clusters remained, including 2 that used machine learning. Extracted data items and grading of results on approaches are provided in Multimedia Appendix 4 [21,23,27-30,32-89]. The specifications of the reported accuracy, reliability, and time efficiency are described in Table 1.

Table 1. Review of the accuracy, test–retest reliability, and time efficiency for automated and machine learning audiometry approaches (2012-2021; N=27 approach clusters).

Type of transducer	Accuracy		Reliability (test–retest)		Time efficiency	
	Reported finding	Values, n (%)	Reported finding	Values, n (%)	Reported finding	Values, n (%)
Air conduction (n=23 approach clusters)						
	RMSD[a]<6 dB[b]	4 (17)	RMSD<6 dB	4 (17)	Acceptable testing time per (partial) audiogram	10 (43)
	RMSD<10 dB	7 (30)	RMSD<10 dB	1 (4)	Acceptable testing time and number of trials per audiogram	2 (9)
	Statistical equivalence	9 (39)	Statistical equivalence	9 (39)	Acceptable testing time and number of trials per frequency	1 (4)
	No statistical equivalence	3 (13)	Not reported	9 (39)	Testing time potential burden	1 (4)
	N/A[c]	N/A	N/A	N/A	Not reported	9 (39)
Bone conduction (n=1 approach cluster)						
	Statistical equivalence	1 (100)	Test–retest not reported	1 (100)	Not reported	1 (100)
Both air and bone conduction (n=3 approach clusters)						
Air conduction						
	RMSD<6 dB	2 (67)	RMSD<6 dB	1 (33)	Acceptable testing time per audiogram	2 (67)
	RMSD<10 dB	1 (33)	RMSD<10 dB	2 (67)	N/A	N/A
Bone conduction						
	RMSD<10 dB	1 (33)	RMSD<6 dB	1 (33)	N/A	N/A
	Statistical equivalence	2 (67)	Test–retest not reported	2 (67)	N/A	N/A
Air and bone conduction						
	N/A	N/A	N/A	N/A	Acceptable testing time per audiogram	1 (33)

[a]RMSD: root mean square deviation.

[b]dB: decibels.

[c]N/A: not applicable.

Accuracy

Accuracy is represented as a comparison against the gold standard or reasonable standard. Most of the automated techniques (14/27, 52%) expressed accuracy in RMSD. Other types of analyses used average differences and SD (10/27, 37%), average thresholds and SD (1/27, 4%) [32], linear regression and correlation coefficients (1/27, 4%) [33], and analysis of variance (1/27, 4%) [34]. The types of analysis used can be seen in Multimedia Appendix 5 [23,32-37,39,40,43,45,48-50,57-59,65,67,68,70,74,77,81,83-85].

Test–Retest Reliability

Test–retest reliability was reported for some automated and machine learning audiometry approaches. Of the 27 approaches, 17 (63%) did not report on test–retest reliability, and 7 (26%) expressed it in RMSD. Other statistical methods used were average differences and SD (6/27, 22%), Pearson product moment correlation coefficients (2/27, 7%) [35,36], standard of variance (1/27, 4%) [37], and repeated analysis of variance (1/27, 4%) [34].

Test Efficiency

Of the 27 approaches, 17 (63%) reported a measure for test efficiency based on the test duration. Test efficiency expressed

in testing time seems to be a standard metric, similar across studies and defined as the time from presenting the first stimulus until the final response of the participant, expressed in seconds or minutes. However, there were disagreement among reports on what to include in the measurement and what groups to use as a reference. Reported time-efficiency measures included the recorded time per frequency, recorded time per unilateral or bilateral air conduction audiogram (between 2 and 7 frequencies) in normal hearing or people with hearing impairment, or full air and bone conduction audiograms in people with hearing impairment. Of the 27 approach clusters, 13 (48%) approach clusters reported acceptable testing times; 3 (11%) approach clusters indicated the number of trials in addition to the testing time for either a bilaterally masked air audiogram [29], unilateral air audiogram [23], or per frequency [38]; 1 (4%) approach cluster that applied Bekesy tracking reported the testing time but was not in the acceptable range [39]; and 10 (37%) approach clusters did not report anything about the testing time.

Test Parameters and Specifications

All tests were self-administered from the point at which the test started. Approximately 15% (4/27) of approaches had the option of switching to a manual audiometry mode. Table 2 summarizes an overview of the test parameters and specifications of the 27 approach clusters, and Table 3 highlights the key contributions. Most of the approaches used adaptive procedures that relied only on the previous response (here referred to as partially adaptive procedures).

The most common example was the (modified) Hughson-Westlake staircase procedure (20/27, 74%), which is based on the classical method of limits [91]. Other partially adaptive procedures applied the method of adjustment, such as the Bekesy tracking method [39] or the *coarse-to-fine focus* algorithm [40]. There was a single report of an approach that did not define the threshold-seeking method but had a built-in protocol to alternate between ears during testing [35]. In contrast, fully adaptive procedures used a complete set of all previous responses. Examples include Bayesian active learning procedures (also referred to as machine learning audiometry; 2/27, 7%) [21,23] and maximum likelihood estimation (2/27, 7%) [37,38]. All machine learning audiometry methods applied active Bayesian model selection, which is a type of shallow machine learning that uses individual models. They apply supervised learning, as every data point is labeled by the participant [22].

Most of the approaches (20/27, 74%) used conventional calibration according to the International Organization for Standardization standards. Of the 27 approaches, 6 (22%) used an unconventional calibration technique. Patel et al [32] determined a reference equivalent threshold level for air conduction for a specific phone–headphone combination using manual audiometry as a reference. Masalski et al [41] used reference levels for calibration for smartphone and transducer combinations, collected under uncontrolled conditions in people with normal hearing. Other calibration techniques set the volume of the device to 50% [42], comparing and adjusting the output level to the input using a sound level meter [34,43], or using Thévenin-equivalent probe calibration [39].

Of the 27 approaches, 22 (82%) were validated in people with normal hearing and hearing impairment. Approximately 7% (4/56) of studies were performed in people with normal hearing [34,36,38]. One of the approach clusters was only validated in a population with hearing impairments using hearing aids as transducers [40]. Automated audiometry was applied across a range of populations. All approaches were applied to adults, except in the study by Patel et al [32] that only included children. Approximately 30% (8/27) approaches were validated in children, including 50% (4/8) of approaches that designed a child-friendly user interface [32,44-46]. Other test populations were older people [47], veterans [48], and persons exposed to occupational noise [49] or ototoxic substances [50]. Automated audiometry has also been applied as an alternative to traditional manual audiometry in low-resource environments [51-53]. The user interface plays an important role in making self-testing feasible in all populations and may require an iterative design process (including clinical pilot studies) [52,54].

Table 2. Description of test parameters and specifications for automated audiometry approaches (2012-2021; N=27).

Test parameters and specifications	Descriptions of approach clusters, n (%)
Threshold-seeking method (underlying algorithm to determine the thresholds)	
Hughson-Westlake (modified)	20 (74)
Machine learning	2 (7)
Bekesy tracking	1 (4)
Other method	4 (15)
Test range (limits of the frequency that can be tested)	
Clinical frequency range (125 Hz-8000 Hz)	18 (67)
Extended high frequencies range (125 Hz-16,000 Hz)	4 (15)
Reduced frequency range	5 (19)
Test range (limits of intensity that can be tested)	
Intensity range (0-100 dB[a] hearing level)	14 (52)
Reduced intensity range	10 (37)
Intensity range not reported	3 (11)
Masking (needed to prevent responses from the nontest ear and obtain the true threshold of the test ear)	
Automated masking	9 (33)
Manual masking	1 (4)
No masking	13 (48)
Masking not reported	4 (15)
Response method (method of recording participants' responses to test stimuli)	
Forced choice	9 (33)
Single response	13 (48)
Forced choice and single response	3 (11)
Not reported	2 (7)
Transducers (method of presenting stimuli, eg, insert phone or supra- or circumaural headphones)	
Air conduction transducers	23 (85)
Air and bone conduction transducers	3 (11)
Only bone conduction transducer	1 (4)
Calibration (unconventional calibration methods are explained in the text)	
Conventional calibration	20 (74)
Unconventional calibration	6 (22)
Calibration not reported	1 (4)
Digital devices (reported hardware needed to run the test)	
Portable audiometer	2 (7)
Computer based	9 (33)
Web-based (requires connectivity)	1 (4)
Smartphone- or tablet-based	1 (4)
Quality control measures (indicators of the reliability of the test)	
Detect false responses	5 (19)
Have noise control	6 (22)
Detect false responses and have noise control	7 (26)
Quality control measures not reported	9 (33)
Validation (highest level of validation reported for each approach cluster)	

Test parameters and specifications	Descriptions of approach clusters, n (%)
Gold standard	22 (82)
Reasonable standard	4 (15)
Proof of concept	1 (4)
Test population (hearing status)	
Normal hearing only	3 (11)
Hearing loss only	1 (4)
Normal hearing and hearing loss	23 (85)
Test population (age)	
Adults only	17 (63)
Children only	1 (4)
Adults and children	9 (33)

[a]dB: decibels

Table 3. Key contributions of the automated and machine learning approaches to the audiological field.

Approach cluster (lead author of first report, reports)	Approach cluster (name)	Key contributions to the field
Bean et al [55]	OtoKiosk	It has the potential to be used in test environments such as examination rooms as a clinical tool for identifying hearing loss via air conduction separating people with normal and impaired hearing.
Chen et al [40]	SHSA[a]	It is a hearing test that runs on a hearing aid, which has statistical equivalence to manual audiometry.
Colsman et al [36]	__[b]	Portable devices that use calibrated headphones result in much higher accuracies than uncalibrated devices.
Corry et al [34]	—	The reliability of audiometer apps should not be assumed. Issues of accuracy and calibration of consumer headphones need to be addressed before such combinations can be used with confidence.
Dewyer et al [33]	Earbone	It is a proof of concept for smartphone-based bone conduction threshold testing.
Foulad et al [43,51,56]	Eartrumpet	It is an iOS-based software app for automated pure-tone hearing testing without the need for additional specialized equipment, yielding hearing test results that approach those of conventional audiometry.
Jacobs et al [50,57]	Oto-ID	They are automated (remote) hearing tests to provide clinicians information for ototoxicity monitoring.
Kung et al [45]	Kids Hearing Game	It includes tablet-based audiometry using game design elements that can be used to test and screen for hearing loss in children who may not have adequate access to resources for a traditional hearing screening.
Liu et al [58]	—	A self-testing system comprising a notebook computer, sound card, and insert earphones is a valid, portable, and sensitive instrument for hearing thresholds self-assessment.
Manganella et al [35]	Agilis	It is an application that detects increased levels of ambient noise when it is programmed to stop the testing.
Margolis et al [30,46,59-61]	AMTAS[c]	AMTAS is designed to fit into the clinical care pathway, including air and bone conduction, and incorporates a quality assessment method (QUALIND) that predicts the accuracy of the test.
Margolis et al [48,62,63]	Home Hearing Test	It is developed and well-suited to provide increased access to hearing testing and support home telehealth programs.
Masalski and Krecicki [41,64,65]	—	It is an automated method that uses smartphone model–specific reference sound levels for calibration in the app. Biological reference sound levels were collected in uncontrolled conditions in people with normal hearing.
Meinke et al [66,67]	WHATS[d]	WHATS is a mobile wireless automated hearing test system in occupational audiometry for obtaining hearing thresholds in diverse test locations without the use of a sound booth.
Patel et al [32]	HearTest[e]	It is a novel, subjective, test-based approach used to calibrate a smartphone–earphone combination with respect to the reference audiometer.
Poling et al [39]	—	Specific Bekesy tracking patterns were identified in people who experienced difficulty converging to a reliable threshold.
Schlittenlacher et al [23]	—	Bayesian active learning methods provide an accurate estimate of hearing thresholds in a continuous range of frequencies.
Schmidt et al [37]	—	A user-operated, 2-alternative, forced choice in combination with the method of maximum likelihood does not require specific operating skills; repeatability is acceptable and is similar to conventional audiometry.
Song et al [21,29,68,69]	MLAG[f]	MLAG is a Bayesian active learning method that determines the most informative next tone, leading to a fast audiogram procedure and threshold estimation in a continuous range of frequencies, with the potential to measure additional variables efficiently.
Sun et al [70]	—	It is an active noise control technology to measure outside the sound booth.
Swanepoel et al [27,47,53,71-75]	KUDUwave	It is an automated portable diagnostic audiometer using improved passive attenuation and real-time environmental noise monitoring, making audiometry possible in unconventional settings.
Swanepoel et al [28,52,54,76-80]	HearTest[g]	It is a smartphone-based automated hearing test applicable in low-resource environments.
Szudek et al [42,81,82]	Uhear	It is an approach that is applicable to the initial evaluation of patients with sudden sensorineural hearing loss before a standard audiogram is available.
Van Tasell and Folkeard [83]	—	Method of adjustment and the Hughson–Westlake method embedded in automated audiometry can be considered equivalent in accuracy to conventional audiometry.

Approach cluster (lead author of first report, reports)	Approach cluster (name)	Key contributions to the field
Vinay et al [38,49]	NEWT[h]	NEWT, which is incorporated inside an active communication earplug, serves as a reliable and efficient method of measuring auditory thresholds, especially in the presence of high background noise.
Whitton et al [84]	—	It is a proof-of-concept study of several self-administered, automated hearing measurements at home, showing statistical equivalency to conventional audiometry in the clinic.
Yeung et al [44,85-89]	Shoebox	It is a method for threshold hearing assessments outside conventional sound booths and with an interface suitable for children.

[a]SHSA: smartphone-based hearing self-assessment.

[b]Not available.

[c]AMTAS: Automated Method for Testing Auditory Sensitivity.

[d]WHATS: Wireless Automated Hearing Test System.

[e]Smartphone-based hearing test app (not yet commercialized).

[f]MLAG: Machine Learning Audiogram.

[g]Automated hearing test commercialized by the hearX group.

[h]NEWT: The New Early Warning Test.

Discussion

Principal Findings

In 2013, evidence for automated audiometry demonstrated similar reliability and accuracy as that of manual audiometry. However, especially for children and bone conduction, the number of reports was limited [24]. In less than a decade, 22 novel approaches and developments across 5 existing approaches had appeared in 56 publications, adding to the 29 papers published before 2013. Promising new developments include the use of machine learning techniques for more time-efficient hearing assessment (2/27, 7%), use of tablets or smartphones as audiometer interface (15/27, 56%), and child-friendly user interfaces (4/27, 15%), including game design elements. The number of approaches that include bone conduction is still limited (4/27, 15%)—only 7% (2/29) more approaches were reported compared with the number reported in 2013 [24].

Accuracy

The required accuracy, reliability, and efficiency depend on the clinical aims and consequences. The ultimate aim of the automated hearing assessment is to deliver clinically actionable estimates of hearing status (ie, the clinician or patient acts appropriately for treatment, given the diagnostic test results). In fully adaptive procedures, the level of precision and confidence needed to conclude the assessment can be set to any level by choosing the proper termination criteria, resulting in different trade-offs. A study by Schmidt et al [37], for instance, aimed for high accuracy and reliability, whereas a study by Heisey et al [29] aimed for high efficiency with machine learning audiometry. Overall, a shift in the type of analysis to demonstrate the accuracy has been observed. In this review, the 2 major types of analysis included were RMSD (14/27, 52%) and average differences and SD (10/27, 37%). In the report by Mahomed et al [24], accuracy was primarily expressed in average differences (11/27, 41%) or thresholds and SD (11/27, 41%). In our view, RMSD is the preferred indicator for accuracy as it has clinical relevance [31], assuming it has already been demonstrated that there is no bias between the automated and manually determined hearing thresholds (eg, signed differences). In traditional clinical terms, automation is equal in accuracy to manual audiometry if the difference is within 6 dB RMSD. Of the 27 automated approaches, 6 (22%) meet this strict accuracy criterion. However, for many applications, the less strict 10 dB RMSD criterium is sufficient, which was achieved by 26% (7/27) additional automated approaches.

For bone conduction measurements, the accuracy was inherently lower than that of air conduction measurements because of conductor placement [30]. However, this reduced accuracy is typically sufficient to address the clinical question of whether conductive or mixed hearing loss is present, as well as choose and evaluate appropriate treatment. The technical feasibility of bone conduction assessments outside of a clinical setting (sound booth) remains difficult. Alternatively, this clinical question can be addressed with other tests, including tympanometry, otoscopy, or a combination of air conduction thresholds for tone and speech stimuli [90]. At least 13 automated techniques had accuracy comparable with that of traditional manual air conduction audiometry, as expressed in RMSD.

A limitation to the impact of achieved test accuracy is the high variation in the interpretation of audiograms by clinicians, regardless of whether those audiograms are determined using an automated or manual approach [92]. Automation can assist clinicians and patients in interpreting the measurement by data-driven automated reporting of accuracy and reliability (including signaling for suspicious outcomes) such as QUALIND [60] or by automated classification for diagnostic purposes (including the type and degree of hearing loss). Examples of automated classification include AMCLASS [93], Autoaudio [94], and data-driven audiogram classification [95].

Reliability

RMSD is also increasingly used as a measure of test–retest reliability. Of the 27 approaches that reported test–retest reliability, 8 (30%) used RMSD as a measure, whereas in 2013, this was only used in 2 (2/29, 7%) studies. Furthermore, 41% (11/27) of approaches did not report on test–retest reliability or used a measure of statistical equivalence that did not allow us

to assess the accuracy. Advances in automated audiometry that increase reliability include procedures to identify invalid responses (5/27, 19%), monitoring environmental noise (6/27, 22%), or both (7/27, 26%) to warn for invalid test conditions, making these tests applicable in more populations and environments. The reliability can be increased, for instance, by alternative response methods, including the forced-choice paradigm [37], or by using machine learning to account for lapses of attention [23]. Digital (health) technologies, including smartphones and tablets, lend themselves to quality control measures for increased reliability with the host of integrated sensors [6].

Efficiency

A fair indicator of efficiency is the overall time required to conduct a test. Most approaches (20/27, 74%) used the modified Hughson–Westlake procedure, of which some (7/20, 35%) showed a similar test duration to manual audiometry. Maximum likelihood procedures demonstrated a 45% reduction in test time in people with normal hearing [38]. Bayesian active learning methods can be extended by adding variables that share some interrelationships using a conjoint estimator that exploits nonlinear interactions between the variables [96]. The resulting machine learning–based automated procedures demonstrated a 30% to 70% reduction in test time compared with manual audiometry for air conduction audiograms in people with normal hearing and hearing impairment [29]. No machine learning approaches had incorporated bone conduction. Therefore, time-efficiency gains compared with full audiogram procedures are not available; however, one can assume that these will yield similar time-efficiency gains. Another indicator of test efficiency is the number of stimuli required to achieve the desired accuracy. This indicator is helpful in optimizing the threshold-seeking part of the approach. Reporting the equivalent time gains under operational conditions is recommended as this can be readily compared with other efficiency gains, including the reduced traveling time if a visit to the outpatient clinic can be replaced for an at-home test or time savings by automating other parts of the clinical care pathway such as interpretation of the outcome. Other aspects of efficiency beyond time that should be considered are cost reductions when enabling task shifting of professionals or the ability to test outside the sound booth.

Future Developments

To obtain an overall indicator of the technical maturity of an approach, developers should be encouraged to use the technology readiness level (TRL) to report the development phase of a technology. TRLs were initially developed in the aerospace industry to estimate the maturity of technology from basic concepts to flight-proven products [97]. To apply TRLs to automated audiometry, further adjustments can be made to fit the hearing health care sector to the version of biomedical TRLs created by the US Army Medical Research and Materiel Command [98]. For those approaches that are ready for operational use, certification (eg, Conformité Européenne and the United States Food and Drug Administration) can further stimulate clinical adoption and iterative improvements based on clinical feedback. In order to be cost-effective, timely, and responsive, certification for digital self-care approaches may need to be less stringent than those for clinical care. A study by Yeung et al [12] proposed alternative procedures for (fast) certification to keep up with the rapidly developing field of visual eHealth tools. Their recommendations might also be applicable to automated hearing assessments, including a rating by health agencies or nongovernmental organizations (eg, a repository of trusted approaches; see Psyberguide [99] as an example of mental health apps reviewed by experts) or adopting the Clinical Laboratory Improvement Amendments model to ensure that approaches comply with the basic requirements of usability, privacy, and security [12]. Following similar certification procedures in the visual and auditory domains may facilitate diagnosis across medical domains. In addition, standards on minimum quality and consensus on what metadata are needed in health applications to describe the test conditions and facilitate interpretation are currently missing.

Limitations

This scoping review included peer-reviewed reports from widely used and recognized scientific databases. A potential limitation is that some of the commercialized automated approaches may have been developed without peer-reviewed reports. Therefore, some automated approaches could be more mature than previously reported. There is no gold standard for reporting audiometry validation studies, which limits a consistent comparison among approaches. Finally, automated procedures may well be embraced by early adopters first, which could lead to projections on suitability that are overly optimistic for users with poorer digital proficiency.

Conclusions and Recommendations

Since 2013, an increasing number of automated audiometry approaches on digital devices have demonstrated similar accuracy, reliability, and time efficiency as conventional manual audiometry. New developments offer features, versatility, and cost-effectiveness beyond manual audiometry. Fully adaptive procedures, including machine learning techniques, seek hearing thresholds more efficiently. Inexpensive digital devices such as smartphones can be turned into audiometers, increasing accessibility and availability. Higher reliability is achievable by signaling invalid test conditions, and child-friendly user interfaces offer a solution to the hard-to-test population. These approaches can be implemented in the clinical care pathway, remote or virtual hearing health care, community-based services, and occupational health care to address the global need for accessible hearing loss diagnosis.

For successful adoption, standardized measures of accuracy, reliability, and efficiency are needed for comparative purposes. Certification and independent reviews may help prospective users select trustworthy approaches. Further reliability can be achieved by determining which difficult-to-test populations may not be appropriate for automated testing and how to detect and then triage these patients to specialized centers. More user-friendly and failsafe procedures that include remote surveillance and quality control can support automated hearing assessment at scale in specific populations and in concert with diagnostic assessments in other medical domains, including visual health and mental well-being [12,99]. Further contextual information, such as standardized metadata, is needed to help

clinicians interpret the context and limitations of test outcomes. If researchers and clinicians deal carefully with their limitations, automated hearing assessments can be designed such that they form an effective part of service delivery for many people who have or are at risk of hearing loss. Automated audiometry can be part of existing care pathways and also enable new service models, including task shifting to community health workers delivering decentralized care, virtual hearing health care, and over-the-counter or direct-to-consumer hearing aid dispensing.

Acknowledgments

The authors would like to thank On Ying Chan for her help during the development of the search string and Raul Sanchez-Lopez for his comments on the draft versions of the paper. In addition, the authors would like to thank Cris P Lanting and Lucas Mens for their suggestions over the course of this study.

Authors' Contributions

LP and JWW made equal contributions, shared first authorship, conducted the abstract screening, independently extracted information directly relevant to the scoping review question, and drafted the manuscript. LP, JWW, and DWS conceptualized the study. RE and DWS supervised the study, reviewed the results, and edited the manuscript. All authors contributed to the data interpretation.

Conflicts of Interest

DWS has a relationship with the hearX Group (Pty) Ltd, which includes equity, consulting, and potential royalties. DWS holds a patent for smartphone-based audiometry as an inventor.

Multimedia Appendix 1
Search strategy.
[DOCX File , 14 KB - jmir_v24i2e32581_app1.docx]

Multimedia Appendix 2
Template table for grading approaches.
[DOCX File , 17 KB - jmir_v24i2e32581_app2.docx]

Multimedia Appendix 3
Estimated mean testing time for conventional manual bilateral audiometry.
[DOCX File , 31 KB - jmir_v24i2e32581_app3.docx]

Multimedia Appendix 4
Graded approaches.
[DOCX File , 188 KB - jmir_v24i2e32581_app4.docx]

Multimedia Appendix 5
Types of statistical analyses for accuracy and reliability.
[DOCX File , 46 KB - jmir_v24i2e32581_app5.docx]

References

1. Haile LM, Kamenov K, Briant PS, Orji AU, Steinmetz JD, Abdoli A, et al. Hearing loss prevalence and years lived with disability, 1990–2019: findings from the Global Burden of Disease Study 2019. Lancet 2021 Mar;397(10278):996-1009. [doi: 10.1016/s0140-6736(21)00516-x]
2. World report on hearing. World Health Organization. 2021. URL: https://www.who.int/publications-detail-redirect/world-report-on-hearing [accessed 2022-01-13]
3. Kamenov K, Martinez R, Kunjumen T, Chadha S. Ear and hearing care workforce: current status and its implications. Ear Hear 2021 Jan 21;42(2):249-257. [doi: 10.1097/AUD.0000000000001007] [Medline: 33480624]
4. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. N Engl J Med 2019 Apr 04;380(14):1347-1358. [doi: 10.1056/nejmra1814259]
5. Wasmann J, Lanting C, Huinck W, Mylanus E, van der Laak JW, Govaerts P, et al. Computational audiology: new approaches to advance hearing health care in the digital age. Ear Hear 2021;42(6):1499-1507 [FREE Full text] [doi: 10.1097/AUD.0000000000001041] [Medline: 33675587]

6. Taylor K, Staunton H, Lipsmeier F, Nobbs D, Lindemann M. Outcome measures based on digital health technology sensor data: data- and patient-centric approaches. NPJ Digit Med 2020;3:97 [FREE Full text] [doi: 10.1038/s41746-020-0305-8] [Medline: 32715091]
7. Swanepoel DW, De Sousa KC, Smits C, Moore DR. Mobile applications to detect hearing impairment: opportunities and challenges. Bull World Health Organ 2019 Sep 03;97(10):717-718. [doi: 10.2471/blt.18.227728]
8. Diagnostics for better health: considerations for global implementation. World Economic Forum - Insight Report. 2021. URL: http://www3.weforum.org/docs/WEF_Diagnostics_for_Better_Health_Considerations_for_Globa_%20Implementation_2021.pdf [accessed 2022-01-13]
9. Bloem BR, Marks WJ, Silva de Lima AL, Kuijf ML, van Laar T, Jacobs BPF, et al. The Personalized Parkinson Project: examining disease progression through broad biomarkers in early Parkinson's disease. BMC Neurol 2019 Jul 17;19(1):160 [FREE Full text] [doi: 10.1186/s12883-019-1394-3] [Medline: 31315608]
10. Gatsios D, Antonini A, Gentile G, Marcante A, Pellicano C, Macchiusi L, et al. Feasibility and utility of mhealth for the remote monitoring of parkinson disease: ancillary study of the PD_manager randomized controlled trial. JMIR Mhealth Uhealth 2020 Jun 29;8(6):e16414 [FREE Full text] [doi: 10.2196/16414] [Medline: 32442154]
11. Philipsen RH, Sánchez CI, Melendez J, Lew WJ, van Ginneken B. Automated chest X-ray reading for tuberculosis in the Philippines to improve case detection: a cohort study. Int J Tuberc Lung Dis 2019 Jul 01;23(7):805-810. [doi: 10.5588/ijtld.18.0004] [Medline: 31439111]
12. Yeung WK, Dawes P, Pye A, Charalambous A, Neil M, Aslam T, et al. eHealth tools for the self-testing of visual acuity: a scoping review. Npj Digit Med 2019 Aug 22;2(1):1-7. [doi: 10.1038/s41746-019-0154-5]
13. Wilson BS, Tucci DL, Merson MH, O'Donoghue GM. Global hearing health care: new findings and perspectives. Lancet 2017 Dec;390(10111):2503-2515. [doi: 10.1016/s0140-6736(17)31073-5]
14. McDaid D, Park A, Chadha S. Estimating the global costs of hearing loss. Int J Audiol 2021 Mar 16;60(3):162-170. [doi: 10.1080/14992027.2021.1883197] [Medline: 33590787]
15. Tsimpida D, Kontopantelis E, Ashcroft D, Panagioti M. Conceptual Model of Hearing Health Inequalities (HHI Model): a critical interpretive synthesis. Trends Hear 2021;25:23312165211002963 [FREE Full text] [doi: 10.1177/23312165211002963] [Medline: 34049470]
16. Livingston G, Sommerlad A, Orgeta V, Costafreda SG, Huntley J, Ames D, et al. Dementia prevention, intervention, and care. Lancet 2017 Dec;390(10113):2673-2734. [doi: 10.1016/S0140-6736(17)31363-6]
17. Manchaiah V, Eikelboom RH, Bennett RJ, Swanepoel DW. International survey of audiologists during the COVID-19 pandemic: effects on the workplace. Int J Audiol 2021 Jun 17:1-8. [doi: 10.1080/14992027.2021.1937348] [Medline: 34137644]
18. Saunders GH, Roughley A. Audiology in the time of COVID-19: practices and opinions of audiologists in the UK. Int J Audiol 2021 Apr;60(4):255-262. [doi: 10.1080/14992027.2020.1814432] [Medline: 32909474]
19. Eksteen S, Launer S, Kuper H, Eikelboom RH, Bastawrous A, Swanepoel DW. Hearing and vision screening for preschool children using mobile technology, South Africa. Bull World Health Organ 2019 Jun 19;97(10):672-680. [doi: 10.2471/blt.18.227876]
20. Békésy GV. A new audiometer. Acta Oto-Laryngologica 2009 Jul 08;35(5-6):411-422. [doi: 10.3109/00016484709123756]
21. Barbour D, Howard R, Song X, Metzger N, Sukesan K, DiLorenzo J, et al. Online machine learning audiometry. Ear Hear 2019;40(4):918-926 [FREE Full text] [doi: 10.1097/AUD.0000000000000669] [Medline: 30358656]
22. Gardner J, Malkomes G, Garnett R, Weinberger K, Barbour D, Cunningham J. Bayesian active model selection with an application to automated audiometry. NeurIPS Proceedings. 2015. URL: https://proceedings.neurips.cc/paper/2015/hash/d9731321ef4e063ebbee79298fa36f56-Abstract.html [accessed 2022-01-13]
23. Schlittenlacher J, Turner RE, Moore BC. Audiogram estimation using Bayesian active learning. J Acoust Soc Am 2018 Jul;144(1):421-430. [doi: 10.1121/1.5047436] [Medline: 30075695]
24. Mahomed F, Swanepoel D, Eikelboom R, Soer M. Validity of automated threshold audiometry: a systematic review and meta-analysis. Ear Hear 2013;34(6):745-752. [doi: 10.1097/01.aud.0000436255.53747.a4] [Medline: 24165302]
25. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
26. Souza NN, Dhar S, Neely ST, Siegel JH. Comparison of nine methods to estimate ear-canal stimulus levels. J Acoust Soc Am 2014 Oct;136(4):1768-1787 [FREE Full text] [doi: 10.1121/1.4894787] [Medline: 25324079]
27. Swanepoel DW, Mngemane S, Molemong S, Mkwanazi H, Tutshini S. Hearing assessment-reliability, accuracy, and efficiency of automated audiometry. Telemed J E Health 2010 Jun;16(5):557-563. [doi: 10.1089/tmj.2009.0143] [Medline: 20575723]
28. van Tonder J, Swanepoel DW, Mahomed-Asmail F, Myburgh H, Eikelboom RH. Automated smartphone threshold audiometry: validity and time efficiency. J Am Acad Audiol 2017 Mar 26;28(3):200-208. [doi: 10.3766/jaaa.16002] [Medline: 28277211]
29. Heisey K, Walker A, Xie K, Abrams J, Barbour D. Dynamically masked audiograms with machine learning audiometry. Ear Hear 2020;41(6):1692-1702 [FREE Full text] [doi: 10.1097/AUD.0000000000000891] [Medline: 33136643]

30. Margolis RH, Glasberg BR, Creeke S, Moore BC. AMTAS: automated method for testing auditory sensitivity: validation studies. Int J Audiol 2010 Mar;49(3):185-194. [doi: 10.3109/14992020903092608] [Medline: 20109081]
31. Guidelines for manual pure-tone threshold audiometry. American Speech-Language-Hearing Association. 2005. URL: https://www.asha.org/policy/gl2005-00014/ [accessed 2022-01-14]
32. Patel K, Thibodeau L, McCullough D, Freeman E, Panahi I. Development and pilot testing of smartphone-based hearing test application. Int J Environ Res Public Health 2021 May 21;18(11) [FREE Full text] [doi: 10.3390/ijerph18115529] [Medline: 34064080]
33. Dewyer NA, Jiradejvong P, Lee DS, Kemmer JD, Sabes JH, Limb CJ. Automated smartphone audiometry: a preliminary validation of a bone-conduction threshold test app. Ann Otol Rhinol Laryngol 2019 Jun 11;128(6):508-515. [doi: 10.1177/0003489419828770] [Medline: 30744390]
34. Corry M, Sanders M, Searchfield G. The accuracy and reliability of an app-based audiometer using consumer headphones: pure tone audiometry in a normal hearing group. Int J Audiol 2017 Sep;56(9):706-710. [doi: 10.1080/14992027.2017.1321791] [Medline: 28485673]
35. Manganella JL, Stiles DJ, Kawai K, Barrett DL, O'Brien LB, Kenna MA. Validation of a portable hearing assessment tool: Agilis Health Mobile Audiogram. Int J Pediatr Otorhinolaryngol 2018 Oct;113:94-98. [doi: 10.1016/j.ijporl.2018.04.010] [Medline: 30174018]
36. Colsman A, Supp GG, Neumann J, Schneider TR. Evaluation of accuracy and reliability of a mobile screening audiometer in normal hearing adults. Front Psychol 2020 Apr 29;11:744 [FREE Full text] [doi: 10.3389/fpsyg.2020.00744] [Medline: 32411050]
37. Schmidt JH, Brandt C, Pedersen ER, Christensen-Dalsgaard J, Andersen T, Poulsen T, et al. A user-operated audiometry method based on the maximum likelihood principle and the two-alternative forced-choice paradigm. Int J Audiol 2014 Jun 10;53(6):383-391. [doi: 10.3109/14992027.2013.879339] [Medline: 24512274]
38. Vinay SN, Svensson UP, Kvaløy O, Berg T. A comparison of test–retest variability and time efficiency of auditory thresholds measured with pure tone audiometry and new early warning test. Appl Acoust 2015 Apr;90:153-159. [doi: 10.1016/j.apacoust.2014.11.002]
39. Poling G, Kunnel T, Dhar S. Comparing the accuracy and speed of manual and tracking methods of measuring hearing thresholds. Ear Hear 2016;37(5):336-340 [FREE Full text] [doi: 10.1097/AUD.0000000000000317] [Medline: 27232075]
40. Chen F, Wang S, Li J, Tan H, Jia W, Wang Z. Smartphone-based hearing self-assessment system using hearing aids with fast audiometry method. IEEE Trans Biomed Circuits Syst 2019:170-179. [doi: 10.1109/tbcas.2018.2878341]
41. Masalski M, Kipiński L, Grysiński T, Kręcicki T. Hearing tests on mobile devices: evaluation of the reference sound level by means of biological calibration. J Med Internet Res 2016 May 30;18(5):e130 [FREE Full text] [doi: 10.2196/jmir.4987] [Medline: 27241793]
42. Szudek J, Ostevik A, Dziegielewski P, Robinson-Anagor J, Gomaa N, Hodgetts B, et al. Can Uhear me now? Validation of an iPod-based hearing loss screening test. J Otolaryngol Head Neck Surg 2012 Apr;41 Suppl 1:78-84. [Medline: 22569055]
43. Foulad A, Bui P, Djalilian H. Automated audiometry using apple iOS-based application technology. Otolaryngol Head Neck Surg 2013 Nov 20;149(5):700-706. [doi: 10.1177/0194599813501461] [Medline: 23963611]
44. Yeung JC, Heley S, Beauregard Y, Champagne S, Bromwich MA. Self-administered hearing loss screening using an interactive, tablet play audiometer with ear bud headphones. Int J Pediatr Otorhinolaryngol 2015 Aug;79(8):1248-1252. [doi: 10.1016/j.ijporl.2015.05.021] [Medline: 26055197]
45. Kung B, Kunda L, Groff S, Miele E, Loyd M, Carpenter DM. Validation Study of Kids Hearing Game: A Self-Administered Pediatric Audiology Application. Perm J 2021 May;25. [doi: 10.7812/TPP/20.157] [Medline: 33970081]
46. Margolis R, Frisina R, Walton J. AMTAS(®): automated method for testing auditory sensitivity: II. air conduction audiograms in children and adults. Int J Audiol 2011 Jul;50(7):434-439. [doi: 10.3109/14992027.2011.553206] [Medline: 21417674]
47. Maclennan-Smith F, Swanepoel DW, Hall JW. Validity of diagnostic pure-tone audiometry without a sound-treated environment in older adults. Int J Audiol 2013 Feb 11;52(2):66-73. [doi: 10.3109/14992027.2012.736692] [Medline: 23140522]
48. Margolis RH, Killion MC, Bratt GW, Saly GL. Validation of the Home Hearing Test™. J Am Acad Audiol 2016 May 06;27(5):416-420. [doi: 10.3766/jaaa.15102] [Medline: 27179261]
49. Henriksen V, Kvaløy O, Svensson UP. Development and calibration of a new automated method to measure air conduction auditory thresholds using an active earplug. Acta Acust United Acust 2014 Jan 01;100(1):113-117. [doi: 10.3813/aaa.918691]
50. Jacobs PG, Silaski G, Wilmington D, Gordon S, Helt W, McMillan G, et al. Development and evaluation of a portable audiometer for high-frequency screening of hearing loss from ototoxicity in homes/clinics. IEEE Trans Biomed Eng 2012 Nov;59(11):3097-3103. [doi: 10.1109/tbme.2012.2204881]
51. Kelly E, Stadler M, Nelson S, Runge C, Friedland DR. Tablet-based screening for hearing loss: feasibility of testing in nonspecialty locations. Otol Neurotol 2018 Apr;39(4):410-416. [doi: 10.1097/MAO.0000000000001752] [Medline: 29494473]
52. Sandström J, Swanepoel D, Laurent C, Umefjord G, Lundberg T. Accuracy and reliability of smartphone self-test audiometry in community clinics in low income settings: a comparative study. Ann Otol Rhinol Laryngol 2020 Jun;129(6):578-584. [doi: 10.1177/0003489420902162] [Medline: 31965808]

53. Visagie A, Swanepoel DW, Eikelboom RH. Accuracy of remote hearing assessment in a rural community. Telemed J E Health 2015 Nov;21(11):930-937. [doi: 10.1089/tmj.2014.0243] [Medline: 26076335]
54. Sandström J, Swanepoel DW, Myburgh HC, Laurent C. Smartphone threshold audiometry in underserved primary health-care contexts. Int J Audiol 2016 Jan 21;55(4):232-238. [doi: 10.3109/14992027.2015.1124294] [Medline: 26795898]
55. Bean BN, Roberts RA, Picou EM, Angley GP, Edwards AJ. Automated audiometry in quiet and simulated exam room noise for listeners with normal hearing and impaired hearing. J Am Acad Audiol 2021 May 25. [doi: 10.1055/s-0041-1728778] [Medline: 34034339]
56. Saliba J, Al-Reefi M, Carriere JS, Verma N, Provencal C, Rappaport JM. Accuracy of mobile-based audiometry in the evaluation of hearing loss in quiet and noisy environments. Otolaryngol Head Neck Surg 2017 Apr 27;156(4):706-711. [doi: 10.1177/0194599816683663] [Medline: 28025906]
57. Dille MF, Jacobs PG, Gordon SY, Helt WJ, McMillan GP. OtoID: new extended frequency, portable audiometer for ototoxicity monitoring. J Rehabil Res Dev 2013;50(7):997-1006 [FREE Full text] [doi: 10.1682/JRRD.2012.09.0176] [Medline: 24301436]
58. Liu Y, Yang D, Xiong F, Yu L, Ji F, Wang Q. Development and validation of a portable hearing self-testing system based on a notebook personal computer. J Am Acad Audiol 2015 Sep 06;26(8):716-723. [doi: 10.3766/jaaa.14092] [Medline: 26333879]
59. Eikelboom RH, Swanepoel DW, Motakef S, Upson GS. Clinical validation of the AMTAS automated audiometer. Int J Audiol 2013 May 03;52(5):342-349. [doi: 10.3109/14992027.2013.769065] [Medline: 23548148]
60. Margolis RH, Saly GL, Le C, Laurence J. Qualind: a method for assessing the accuracy of automated tests. J Am Acad Audiol 2007 Jan;18(1):78-89. [doi: 10.3766/jaaa.18.1.7] [Medline: 17252960]
61. Margolis R, Moore B. AMTAS(®): automated method for testing auditory sensitivity: III. sensorineural hearing loss and air-bone gaps. Int J Audiol 2011 Jul;50(7):440-447. [doi: 10.3109/14992027.2011.575085] [Medline: 21668325]
62. Margolis RH, Bratt G, Feeney MP, Killion MC, Saly GL. Home hearing test: within-subjects threshold variability. Ear Hear 2018;39(5):906-909 [FREE Full text] [doi: 10.1097/AUD.0000000000000551] [Medline: 29356703]
63. Mosley CL, Langley LM, Davis A, McMahon CM, Tremblay KL. Reliability of the home hearing test: implications for public health. J Am Acad Audiol 2019 Mar;30(3):208-216 [FREE Full text] [doi: 10.3766/jaaa.17092] [Medline: 30461396]
64. Masalski M, Kręcicki T. Self-test web-based pure-tone audiometry: validity evaluation and measurement error analysis. J Med Internet Res 2013;15(4):e71 [FREE Full text] [doi: 10.2196/jmir.2222] [Medline: 23583917]
65. Masalski M, Grysiński T, Kręcicki T. Hearing tests based on biologically calibrated mobile devices: comparison with pure-tone audiometry. JMIR Mhealth Uhealth 2018 Jan 10;6(1):e10 [FREE Full text] [doi: 10.2196/mhealth.7800] [Medline: 29321124]
66. Magro I, Clavier O, Mojica K, Rieke C, Eisen E, Fried D, et al. Reliability of tablet-based hearing testing in nicaraguan schoolchildren: a detailed analysis. Otol Neurotol 2020 Mar;41(3):299-307. [doi: 10.1097/MAO.0000000000002534] [Medline: 31851067]
67. Meinke DK, Norris JA, Flynn BP, Clavier OH. Going wireless and booth-less for hearing testing in industry. Int J Audiol 2017 Dec 15;56(sup1):41-51 [FREE Full text] [doi: 10.1080/14992027.2016.1261189] [Medline: 27976975]
68. Song X, Wallace B, Gardner J, Ledbetter N, Weinberger K, Barbour D. Fast, continuous audiogram estimation using machine learning. Ear Hear 2015;36(6):326-335 [FREE Full text] [doi: 10.1097/AUD.0000000000000186] [Medline: 26258575]
69. Heisey KL, Buchbinder JM, Barbour DL. Concurrent bilateral audiometric inference. Acta Acust United Acust 2018 Sep 01;104(5):762-765. [doi: 10.3813/aaa.919218]
70. Sun C, Liu Y, Wang X. An automated hearing test equipment based on active noise control technology. In: Proceedings of the IEEE International Instrumentation and Measurement Technology Conference (I2MTC). 2019 Presented at: IEEE International Instrumentation and Measurement Technology Conference (I2MTC); May 20-23, 2019; Auckland, New Zealand p. 1-5. [doi: 10.1109/i2mtc.2019.8827137]
71. Brennan-Jones CG, Eikelboom RH, Swanepoel DW, Friedland PL, Atlas MD. Clinical validation of automated audiometry with continuous noise-monitoring in a clinically heterogeneous population outside a sound-treated environment. Int J Audiol 2016 Sep 20;55(9):507-513. [doi: 10.1080/14992027.2016.1178858] [Medline: 27206551]
72. Swanepoel DW, Biagio L. Validity of diagnostic computer-based air and forehead bone conduction audiometry. J Occup Environ Hyg 2011 Apr 22;8(4):210-214. [doi: 10.1080/15459624.2011.559417] [Medline: 21391065]
73. Govender S, Mars M. Validity of automated threshold audiometry in school aged children. Int J Pediatr Otorhinolaryngol 2018 Feb;105:97-102. [doi: 10.1016/j.ijporl.2017.12.008] [Medline: 29447828]
74. Storey KK, Muñoz K, Nelson L, Larsen J, White K. Ambient noise impact on accuracy of automated hearing assessment. Int J Audiol 2014 Oct 09;53(10):730-736. [doi: 10.3109/14992027.2014.920110] [Medline: 24909592]
75. Swanepoel DW, Matthysen C, Eikelboom R, Clark J, Hall JW. Pure-tone audiometry outside a sound booth using earphone attentuation, integrated noise monitoring, and automation. Int J Audiol 2015;54(11):777-785. [doi: 10.3109/14992027.2015.1072647] [Medline: 26514954]
76. Bornman M, Swanepoel D, De Jager LB, Eikelboom R. Extended high-frequency smartphone audiometry: validity and reliability. J Am Acad Audiol 2019 Mar;30(3):217-226. [doi: 10.3766/jaaa.17111] [Medline: 30461416]

77. Swanepoel DW, Myburgh HC, Howe DM, Mahomed F, Eikelboom RH. Smartphone hearing screening with integrated quality control and data management. Int J Audiol 2014 Dec;53(12):841-849. [doi: 10.3109/14992027.2014.920965] [Medline: 24998412]
78. Brittz M, Heinze B, Mahomed-Asmail F, Swanepoel DW, Stoltz A. Monitoring hearing in an infectious disease clinic with mhealth technologies. J Am Acad Audiol 2019 Jun;30(6):482-492. [doi: 10.3766/jaaa.17120] [Medline: 30461403]
79. Corona AP, Ferrite S, Bright T, Polack S. Validity of hearing screening using hearTest smartphone-based audiometry: performance evaluation of different response modes. Int J Audiol 2020 Sep;59(9):666-673. [doi: 10.1080/14992027.2020.1731767] [Medline: 32134341]
80. Rodrigues LC, Ferrite S, Corona AP. Validity of heartest smartphone-based audiometry for hearing screening in workers exposed to noise. J Am Acad Audiol 2021 Feb;32(2):116-121. [doi: 10.1055/s-0040-1718931] [Medline: 33296933]
81. Handzel O, Ben-Ari O, Damian D, Priel MM, Cohen J, Himmelfarb M. Smartphone-based hearing test as an aid in the initial evaluation of unilateral sudden sensorineural hearing loss. Audiol Neurootol 2013;18(4):201-207. [doi: 10.1159/000349913] [Medline: 23689282]
82. Khoza-Shangase K, Kassner L. Automated screening audiometry in the digital age: exploring uhear™ and its use in a resource-stricken developing country. Int J Technol Assess Health Care 2013 Jan;29(1):42-47. [doi: 10.1017/S0266462312000761] [Medline: 23298579]
83. Van Tasell DJ, Folkeard P. Reliability and accuracy of a method of adjustment for self-measurement of auditory thresholds. Otol Neurotol 2013 Jan;34(1):9-15. [doi: 10.1097/MAO.0b013e318278c05d] [Medline: 23202154]
84. Whitton JP, Hancock KE, Shannon JM, Polley DB. Validation of a self-administered audiometry application: an equivalence study. Laryngoscope 2016 Oct 02;126(10):2382-2388. [doi: 10.1002/lary.25988] [Medline: 27140227]
85. Yeung J, Javidnia H, Heley S, Beauregard Y, Champagne S, Bromwich M. The new age of play audiometry: prospective validation testing of an iPad-based play audiometer. J Otolaryngol Head Neck Surg 2013 Mar 11;42(1):21 [FREE Full text] [doi: 10.1186/1916-0216-42-21] [Medline: 23663317]
86. Thompson GP, Sladen DP, Borst BJ, Still OL. Accuracy of a tablet audiometer for measuring behavioral hearing thresholds in a clinical population. Otolaryngol Head Neck Surg 2015 Nov;153(5):838-842. [doi: 10.1177/0194599815593737] [Medline: 26183518]
87. Bastianelli M, Mark AE, McAfee A, Schramm D, Lefrançois R, Bromwich M. Adult validation of a self-administered tablet audiometer. J Otolaryngol Head Neck Surg 2019 Nov 07;48(1):59 [FREE Full text] [doi: 10.1186/s40463-019-0385-0] [Medline: 31699157]
88. Vijayasingam A, Frost E, Wilkins J, Gillen L, Premachandra P, Mclaren K, et al. Tablet and web-based audiometry to screen for hearing loss in adults with cystic fibrosis. Thorax 2020 Aug;75(8):632-639. [doi: 10.1136/thoraxjnl-2019-214177] [Medline: 32409613]
89. Yalamanchali S, Albert RR, Staecker H, Nallani R, Naina P, J Sykes K. Evaluation of portable tablet-based audiometry in a South Indian population. Indian J Otolaryngol Head Neck Surg 2020 Sep 02. [doi: 10.1007/s12070-020-02094-3]
90. De Sousa KC, Swanepoel DW, Moore DR, Myburgh HC, Smits C. Improving sensitivity of the digits-in-noise test using antiphasic stimuli. Ear Hear 2020;41(2):442-450 [FREE Full text] [doi: 10.1097/AUD.0000000000000775] [Medline: 31425362]
91. Gescheider G. Psychophysics: The Fundamentals. East Sussex, United Kingdom: Psychology Press; 2013.
92. Brennan-Jones C, Eikelboom R, Bennett R, Tao K, Swanepoel D. Asynchronous interpretation of manual and automated audiometry: agreement and reliability. J Telemed Telecare 2016 Sep 20;24(1):37-43. [doi: 10.1177/1357633x16669899]
93. Margolis R, Saly G. Asymmetric hearing loss: definition, validation, and prevalence. Otol Neurotol 2008 Jun;29(4):422-431. [doi: 10.1097/MAO.0b013e31816c7c09] [Medline: 18418281]
94. Crowson M, Lee J, Hamour A, Mahmood R, Babier A, Lin V, et al. AutoAudio: deep learning for automatic audiogram interpretation. J Med Syst 2020:163. [doi: 10.1101/2020.04.30.20086637]
95. Charih F, Bromwich M, Mark AE, Lefrançois R, Green JR. Data-driven audiogram classification for mobile audiometry. Sci Rep 2020 Mar 03;10(1):3962 [FREE Full text] [doi: 10.1038/s41598-020-60898-3] [Medline: 32127604]
96. Barbour DL, DiLorenzo JC, Sukesan KA, Song XD, Chen JY, Degen EA, et al. Conjoint psychometric field estimation for bilateral audiometry. Behav Res Methods 2019 Jun;51(3):1271-1285 [FREE Full text] [doi: 10.3758/s13428-018-1062-3] [Medline: 29949072]
97. Héder M. From NASA to EU: the evolution of the TRL scale in public sector innovation. Innov J 2017;22(2):1-23 [FREE Full text]
98. Technology Readiness Assessment (TRA) Deskbook. Department of Defense. 2009. URL: https://www.skatelescope.org/public/2011-11-18_WBS-SOW_Development_Reference_Documents/DoD_TRA_July_2009_Read_Version.pdf [accessed 2022-01-14]
99. Garland A, Jenveja A, Patterson J. Psyberguide: a useful resource for mental health apps in primary care and beyond. Fam Syst Health 2021 Mar;39(1):155-157. [doi: 10.1037/fsh0000587] [Medline: 34014736]

Abbreviations

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RMSD: root mean square deviation
TRL: technology readiness level

Edited by R Kukafka; submitted 03.08.21; peer-reviewed by D Allen, V Singh; comments to author 28.10.21; revised version received 01.12.21; accepted 16.12.21; published 02.02.22.

Please cite as:
Wasmann JW, Pragt L, Eikelboom R, Swanepoel DW
Digital Approaches to Automated and Machine Learning Assessments of Hearing: Scoping Review
J Med Internet Res 2022;24(2):e32581
URL: https://www.jmir.org/2022/2/e32581
doi:10.2196/32581
PMID:34919056

Review

Implementation of eHealth to Support Assessment and Decision-making for Residents With Dementia in Long-term Care: Systematic Review

Juliet Gillam[1], BSc (Hons), MSc; Nathan Davies[2,3], BSc (Hons), MSc, PhD; Jesutofunmi Aworinde[1], BSc (Hons), MSc; Emel Yorganci[1], BSc, MSc; Janet E Anderson[4], BBSc, MPsych, PhD; Catherine Evans[1,5], BSc (Hons), MSc, PhD

[1]Cicely Saunders Institute, King's College London, London, United Kingdom

[2]Centre for Ageing Population Studies, Research Department of Primary Care and Population Health, University College London, London, United Kingdom

[3]Centre for Dementia Palliative Care Research, Marie Curie Palliative Care Research Department, University College London, London, United Kingdom

[4]School of Health Sciences, City, University of London, London, United Kingdom

[5]Sussex Community National Health Service Foundation Trust, Brighton General Hospital, Brighton, United Kingdom

Corresponding Author:
Juliet Gillam, BSc (Hons), MSc
Cicely Saunders Institute
King's College London
Bessemer Road
London, SE5 9PJ
United Kingdom
Phone: 44 7500 708 293
Email: juliet.h.gillam@kcl.ac.uk

Abstract

Background: As dementia progresses, symptoms and concerns increase, causing considerable distress for the person and their caregiver. The integration of care between care homes and health care services is vital to meet increasing care needs and maintain quality of life. However, care home access to high-quality health care is inequitable. eHealth can facilitate this by supporting remote specialist input on care processes, such as clinical assessment and decision-making, and streamlining care on site. How to best implement eHealth in the care home setting is unclear.

Objective: The aim of this review was to identify the key factors that influence the implementation of eHealth for people living with dementia in long-term care.

Methods: A systematic search of Embase, PsycINFO, MEDLINE, and CINAHL was conducted to identify studies published between 2000 and 2020. Studies were eligible if they focused on eHealth interventions to improve treatment and care assessment or decision-making for residents with dementia in care homes. Data were thematically analyzed and deductively mapped onto the 6 constructs of the adapted Consolidated Framework for Implementation Research (CFIR). The results are presented as a narrative synthesis.

Results: A total of 29 studies were included, focusing on a variety of eHealth interventions, including remote video consultations and clinical decision support tools. Key factors that influenced eHealth implementation were identified across all 6 constructs of the CFIR. Most concerned the inner setting construct on requirements for implementation in the care home, such as providing a conducive learning climate, engaged leadership, and sufficient training and resources. A total of 4 novel subconstructs were identified to inform the implementation requirements to meet resident needs and engage end users.

Conclusions: Implementing eHealth in care homes for people with dementia is multifactorial and complex, involving interaction between residents, staff, and organizations. It requires an emphasis on the needs of residents and the engagement of end users in the implementation process. A novel conceptual model of the key factors was developed and translated into 18 practical recommendations on the implementation of eHealth in long-term care to guide implementers or innovators in care homes. Successful implementation of eHealth is required to maximize uptake and drive improvements in integrated health and social care.

(J Med Internet Res 2022;24(2):e29837) doi:10.2196/29837

KEYWORDS

telemedicine; implementation science; dementia; long-term care; systematic review

Introduction

Background

Dementia is a progressive, complex neurodegenerative condition with a multitude of types and clinical presentations. It is the leading cause of death in the United Kingdom [1] and is projected to have the highest global proportional rise (264%) in suffering associated with a need for palliative care [2]. Dementia is characterized by a complexity of care needs, which advance as the condition progresses [3]. These needs span multiple domains of health care, and multimorbidity is common [4]. Symptoms such as pain and breathlessness [5,6] cause significant distress for the person and their caregiver and increase toward the end of life [7].

Over half of the people with dementia (58%) die in care homes in England [8]. They are the main providers of end-of-life dementia care, with the average life expectancy on admission for a resident with dementia being 1 to 2 years [9]. The term *care home* in the United Kingdom refers to both residential and nursing homes. These differ with regard to the provision of input from health care professionals, with nursing homes providing additional access to 24-hour on-site nursing care. Care homes require the resources to deliver multidisciplinary care to meet these advancing and acute dementia-specific needs [10]. Providing access to good quality, continuous care throughout the dementia trajectory is essential. This can be achieved by integrating care homes with primary care, palliative care, and dementia care teams to enable multidisciplinary and specialist input on vital care processes [11].

Care needs change and develop over time and cause considerable distress if left unmet [6]. A comprehensive assessment by a multidisciplinary team to review medical, functional, mental, and social abilities is essential for this population with complex needs [12]. Interprofessional collaboration is also required to share clinical expertise and experience, best inform complex clinical decisions concerning treatment for multimorbidities, and deliver appropriate care [13]. Integrating these processes across services is widely acknowledged to improve person-centered treatment outcomes for older people with complex needs [11] and reduce detrimental transitions between settings occurring in the final years of life, such as unplanned hospitalization [14].

To integrate services and deliver continuous and coordinated care, established methods of communication are required to share information between systems about residents' care needs and outcomes. A way to facilitate this is through the use of eHealth [15], which is defined as "health services and information delivered or enhanced through the internet and related technologies" [16], encompassing an array of interventions that enable care to be delivered remotely.

The COVID-19 pandemic has had major implications for the use of eHealth in care homes. In response, former barriers to the integration of health and social care services in England have been liberalized, with restrictions on information sharing between services relaxed [17]. The pandemic highlighted the lack of systems for efficiently sharing information between services nationally and internationally [18,19]. This compromised the delivery of care to meet the escalating health care needs and residents' quality of life. Innovation was required to communicate between services for the comprehensive assessment of symptoms and management of care and outcomes. eHealth provides a way to do this [20].

Objective

There is little evidence to guide the design and implementation of eHealth resources to manage symptoms and concerns for people with dementia in care homes. The key task at hand is to understand how we can scale-up eHealth interventions to embed them in routine care. Despite positive findings regarding their benefit [21], eHealth interventions are yet to attain widespread implementation in care homes, and the way of effectively achieving this remains unclear. This study aimed to explore factors that influence the implementation of eHealth interventions to support assessment and decision-making regarding care and treatment for people with dementia in care homes.

Methods

Design

The review is reported in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses; Multimedia Appendix 1) guidelines. A systematic review, using narrative synthesis and thematic analysis, was conducted to identify common facilitators of and barriers to eHealth implementation in care homes. The synthesis followed the guidance of Popay et al [22], which provides a specific direction for reviews concerned with the implementation of interventions. The protocol for this review was registered on PROSPERO (International Prospective Register of Systematic Reviews; registration number CRD42020184587). Before registering this review, a search of PROSPERO was conducted to ensure that no similar reviews were underway. Our final search was performed in the week the review was registered (June 1, 2020). Originally, the scope of the review also included collecting data on intervention effectiveness and key components; however, given the breadth of evidence relating to these outcomes, the findings regarding effectiveness will be reported in a second review.

Search Strategy

A total of 4 databases (Embase, PsycINFO, MEDLINE, and CINAHL) were searched for studies published in English from January 2000, with the final search being on April 28, 2020. The year 2000 was chosen as the *cutoff* year to exclude eHealth that may be outdated in the context of today's technological

advancements. A search strategy was developed with the help of an information support specialist and informed by scoping the literature for the types of eHealth interventions currently in use (see Multimedia Appendix 2 for the full search strategy). A combination of Medical Subject Headings terms and keywords was used to develop a strategy based on the following concepts: *dementia* AND *care homes* AND *eHealth* AND *assessment* OR *decision-making*. The search strategy was complemented through reference chaining and citation tracking using Google Scholar, following the initial identification of studies from the database search. The eligibility criteria were developed using the PICO (population, intervention, control, and outcomes) acronym, following recommendations on its suitability and enhanced sensitivity when conducting qualitative systematic reviews [23]. The criteria are outlined in Textbox 1.

Textbox 1. Eligibility criteria.

Inclusion criteria

- Population: residents with a diagnosis of dementia residing in a long-term care facility; studies that included residents with dementia in a mixed population
- Intervention: eHealth interventions that aimed to facilitate comprehensive assessment of care home residents or improve decision-making about care and treatment; eHealth interventions that enable care coordination between practitioners and the sharing of information between settings to facilitate integrated working, such as between care homes and health care services
- Outcome: data relating to factors that influence or inhibit the implementation of eHealth interventions in care homes
- Comparator: no restrictors
- Study design: all study designs that reported data relating to implementation

Exclusion criteria

- Population: people with a diagnosis of dementia living at home or staying in short-term care or acute care settings; studies that did not mention dementia in the population
- Intervention: nondigitalized intervention studies; eHealth interventions that did not focus on aiding comprehensive assessment or supporting clinical decision-making; interventions that monitored clinical signs only, such as a motion sensor, or recorded biodata, for example, blood pressure remote monitoring, were out of scope
- Outcome: not applicable
- Comparator: not applicable
- Study design: opinion pieces

Study Selection Procedure

Studies identified from the search were exported to Endnote [24]. Duplicates were identified and removed. All titles and abstracts were screened by 1 researcher (JG). Approximately 20% were randomly selected for blinded double screening by 2 independent researchers (TA and EY) as a calibration process to test the application of the eligibility criteria. Once an agreement of 90% was confirmed between the reviewers, the eligibility criteria were applied to all identified studies. This was undertaken to screen for eligibility at the title and abstract screening and again at the full paper review. Discrepancies were resolved through discussion.

Quality Appraisal

The quality of publications was appraised using the Critical Appraisal Skills Programme (CASP) tool [25], which was appropriate for the study design. It was conducted by 1 author (JG) and reviewed independently by 2 authors (CE and ND). Where the design was not amenable to the CASP, alternative tools were used, including the Mixed Methods Appraisal Tool (MMAT) [26] and the Joanna Briggs Institute (JBI) critical appraisal tools [27]. No studies were excluded only on the basis of their appraised quality; rather, it was conducted to help understand and describe the studies.

Data Extraction

A standardized data extraction tool was developed in Microsoft Excel and informed by the review questions. Extracted data included study aim, country of origin, design, population of interest, setting, eHealth intervention (type, components, and summary), methods of data collection and analysis, outcomes regarding intervention implementation and conclusions, implications, and limitations. Implementation data were extracted from both the results and discussion sections to capture relevant findings relating to the authors' observations of why an intervention was or was not effective.

Data Analysis and Synthesis

A deductive thematic analytic approach was undertaken, underpinned by the adapted version of the Consolidated Framework for Implementation Research (CFIR) [28]. The original CFIR [29] comprises 5 broad theory-based constructs and 39 subconstructs within these: intervention characteristics, inner setting, outer setting, individual characteristics, and implementation process. It is a comprehensive and practical guide for assessing the potential factors that influence implementation. A recent adaptation of the framework added a sixth construct [28], *patient needs*, acknowledging the importance of person-centered care in health care interventions and the paucity of attention it frequently receives in

implementation frameworks. Given the nature of this review, the adapted CFIR was used.

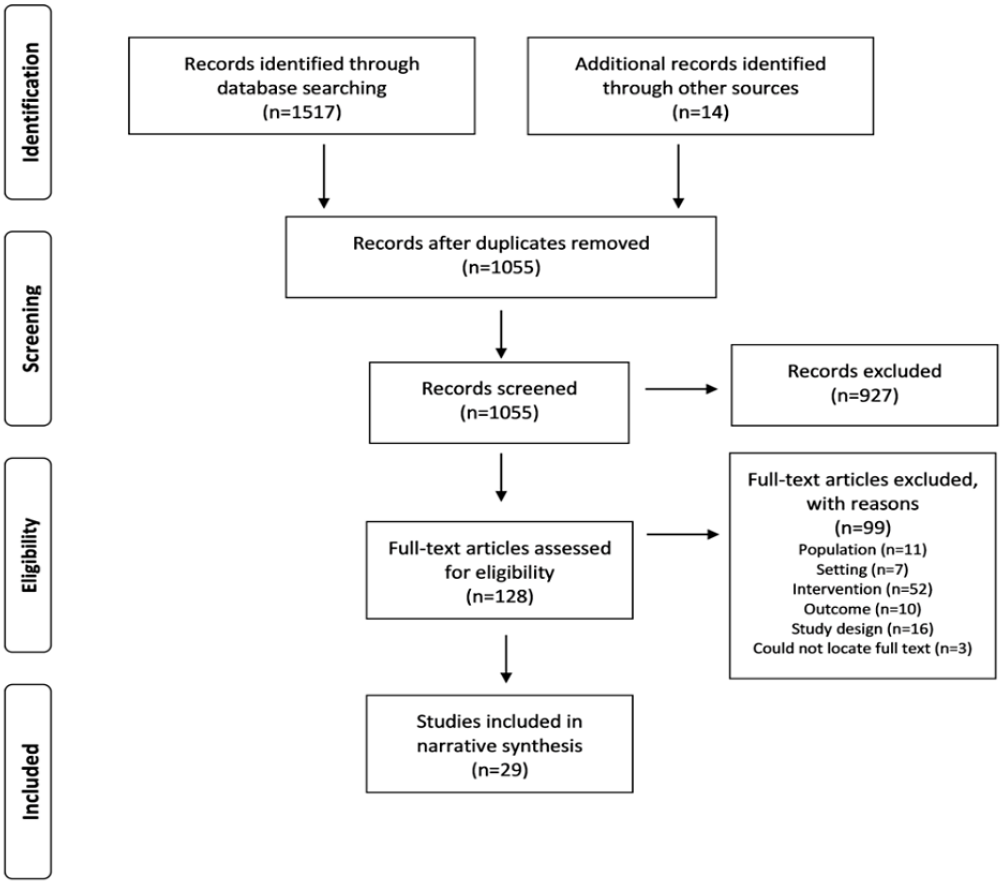
Preliminary synthesis involved tabulation to organize the findings and compare data across different studies. The data from each study were then mapped onto the framework. Common themes that arose across the studies in line with the subconstructs of the framework were then synthesized to form a narrative regarding facilitating elements of and inhibiting barriers to successful implementation. Where data did not align to a subconstruct, an inductive thematic analytic approach was undertaken to avoid biasing data and identify gaps in the framework when applying it to this context. The themes were then developed to identify additional constructs to adapt the framework to this specific context.

Results

Summary

A total of 1055 papers were screened by title and abstract, of which 128 (12.13%) full-text articles were assessed; of the 128 articles, 29 (22.7%) met the eligibility criteria for the review (Figure 1 [30]). The included studies reported 27 unique interventions that aimed to facilitate comprehensive assessment of care home residents or improve decision-making surrounding care and treatment, published between 2000 and 2020. The sample sizes ranged from 5 to 4171 for residents and 6 to 609 for carers. Approximately 31% (9/29) of studies omitted the number of participants.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart [30].



Of the 29 articles, the included studies comprised 7 (24%) randomized controlled trials, 7 (24%) quasi-experimental studies [31-37], 6 (21%) qualitative studies [38-43], 3 (10%) descriptive studies [44-46], 2 (7%) mixed methods studies [47,48], 2 (7%) cohort studies [49,50], and 2 (7%) cross-sectional studies [51,52].

All studies included adults with dementia, focusing specifically on requirements for people with dementia (11/29, 38%) [31,32,38-40,43,47,49,50,53,54] or within a mixed population (18/29, 62%) [33-37,41,42,44-46,48,51,52,55-59]. The average percentage of residents with dementia was 52.8% in the mixed population studies that reported the proportion (6/29, 21%) [35,37,44,46,55,56]. Approximately 41% (12/29) of studies did not delineate the proportion of residents with dementia [33,34,36,41,42,45,48,51,52,57-59] (see Table 1 for a summary of study characteristics).

Table 1. Summary of characteristics of included studies grouped by population residents.

Characteristics, study, and country of origin	Study design	Population (n)	Age (years), mean	Setting (n)	Type of eHealth intervention
Population with dementia and specific requirements					
Catic et al [49], United States[a]	Cohort	Dementia (46)	82.5	Nursing home (11)	Video consultation
Gordon et al [50], United States[a]	Matched-cohort study	Dementia (115)	Not specified	Nursing home (11)	Video consultation
Klein et al [38], Australia	Qualitative	Dementia (5)	59-88	Regional aged care facility (1)	PDA
Lee et al [53], Korea	Quasi-experimental	Dementia (53)	Not specified	Nursing home (1)	Video consultations and computerized system
Lyketsos et al [54], United States	Quasi-experimental	Dementia (not specified)	Not specified	Long-term dementia facility (1)	Video consultations
Mitchell et al [31], United States	Cluster RCT[b]	Dementia (402)	86.7	Nursing home (64)	Video decision support tool
Qadri et al [47], United States	Mixed methods	Dementia (not specified)	Not specified	Nursing home (3)	PDA
Moniz-Cook et al [32], United Kingdom	Cluster RCT	Dementia (832)	Not specified	Care home (58)	Computerized decision support tool
Piau et al [39], France	Qualitative	Dementia (90)	Not specified	Long-term care facility (10)	Video consultations
Shiells et al [40], Czechia	Qualitative	Dementia (not specified)	Not specified	Nursing home (3)	Electronic patient records
Keenan et al [43], United Kingdom	Qualitative	Dementia (not specified)	Not specified	Care home (27)	Computerized decision support tool
Population with mixed requirements (dementia and nondementia)					
Salles et al [45], France	Descriptive	Mixed (304); others: wounds and psychiatric	85.6	Nursing home (1)	Video consultations
Daly et al [33], United States	RCT	Mixed (22)	86	Long-term care facility (1)	Electronic patient records
De Luca et al [34], Italy	RCT	Mixed (59)	79.1	Nursing home (1)	Video consultations and tele-counseling
Johnston and Jones [44], United States	Descriptive	Mixed; dementia (52.5%); others: delirium, depression, and dysthymia	79.3	Rural skilled nursing facility (1)	Video consultations
Krüger et al [55], Norway	Quasi-experimental	Mixed; dementia (76.6%) and stroke (23.4%)	84.4	Nursing home (7)	Electronic patient records with decision support tool
Mitchell et al [35], United States	Cluster RCT	Mixed; dementia (70%)	Not specified	Nursing home (intervention=119; control=241)	Video decision support tool
Perri et al [56], Canada	Quasi-experimental	Mixed; dementia (69%); other: cardiovascular, respiratory, frailty, and psychiatric.	87	Long-term care facility (1)	Video consultations
Alexander [46], United States[c]	Descriptive	Mixed; dementia (20), Alzheimer (13), and other: osteoarthritis, pneumonia, and cerebrovascular accident	Not specified	Nursing home (3)	PDA, electronic patient records, and decision support tool

Characteristics, study, and country of origin	Study design	Population (n)	Age (years), mean	Setting (n)	Type of eHealth intervention
Mor et al [37], United States	RCT	Mixed (not specified); dementia (30%); others: cardiopulmonary	Not specified	Nursing home (360)	Video decision support tool
Vuorinen [41], New Zealand	Qualitative	Mixed (not specified)	≥65 (intervention)	Long-term care facility dementia unit (1)	Web-based assessment tool
Weiner et al [36], United States	RCT	Mixed (369)	64	Nursing home (1)	Video consultations
Pillemer et al [57], United States	Quasi-experimental	Mixed (761)	79.4	Nursing home (10)	Electronic patient records and PDAs
O'Mahony et al [58], United States	Quasi-experimental	Mixed (not specified); dementia, cancer, chronic obstructive pulmonary disease, liver disease, and renal failure	Not specified	Skilled nursing facilities (2)	Video consultations
Munyisia et al [48], Australia	Mixed methods	Mixed (not specified)	Not specified	Nursing home (1) and specialized home (1)	Electronic patient records
Alexander et al [42], United States[c]	Qualitative	Mixed (not specified)	Not specified	Nursing home (4)	PDA, electronic patient records, and decision support tool
Bjarnadottir et al [51], United States	Cross-sectional analysis	Mixed (not specified)	Not specified	Nursing home (927)	Electronic patient records
Wakefield et al [52], United States	Cross-sectional analysis/longitudinal	Mixed (62); dementia, seizure, Parkinson, and urinary tract infections	72	Long-term care facility (1)	Video consultations
Fossum et al [59], Sweden	Quasi-experimental	Mixed (491)	84.5	Nursing home (15)	Electronic patient records with decision support tool

[a]Articles from the same study.

[b]RCT: randomized controlled trial.

[c]Articles from the same study.

Quality Appraisal

The included studies varied in their quality. In general, the CASP criteria for the experimental studies identified consistent reporting of clear and focused aims, with appropriate methodologies to address the research questions. Weaknesses were related to small sample size, evidence of selection and attrition bias, poor description of analysis, and nonblinding. Descriptive studies were generally well-reported (average 86% of MMAT criteria met) and quasi-experimental studies (average 77.7% of JBI criteria met). The weaker study design used mixed methods (average 50% of MMAT criteria met) and cross-sectional studies (average 56.2% of JBI criteria met). Weaknesses pertained to the management of confounding factors and the integration of qualitative and quantitative data. Multimedia Appendix 3 [31-59], details the quality appraisal results for each study using the respective appraisal checklist.

CFIR Constructs Associated With eHealth Implementation

Table 2 details the respective CFIR constructs and subconstructs that were identified as important determinants for implementation. The number of subconstructs identified per study ranged from 1 to 13, with a median of 6. No major differences in implementation requirements were identified between studies with a specific focus on requirements for people with dementia and studies reporting on dementia within a mixed population (Multimedia Appendix 4). Findings for each CFIR construct and respective subconstructs is presented in turn. Constructs that were identified in ≤2 studies are not presented as a narrative because of insufficient data.

Table 2. Factors identified to influence the implementation of eHealth interventions in a care home (N=29).

CFIR[a] construct and subconstructs	Definition in the context of care homes and integrated care for people with dementia	Total, n (%)
Intervention characteristics: aspects of eHealth that might affect implementation success		
Intervention source	How end users perceive the legitimacy of the eHealth source—whether it has been developed internally as a response to a problem in the care home or externally	0
Evidence strength and quality	Stakeholder perception of the strength of the evidence supporting the belief that eHealth will produce the desired outcomes from sources such as published literature	0
Relative advantage [36-39,47,48,54-56]	Whether stakeholders perceive eHealth as advantageous over current practice	9 (31)
Adaptability [37,40,48,49,52,55,59]	How interoperable eHealth is with current care home information technology systems	7 (24)
Trialability	Whether eHealth can be tested initially on a small scale, such as piloted in a small number of care homes	0
Complexity [36-38,40,41,44,45,47,48,53,54,56,58]	How simple and user-friendly end users perceive eHealth to be within routine care	13 (45)
Design quality and packaging	Stakeholder perception of the physical presentation of the eHealth intervention	0
Cost [36,37,47,49-51,53,54]	Cost associated with implementing eHealth	8 (28)
Patient needs: the extent to which resident needs are known and prioritized by the care home		
Clinical benefit[b] [34,36,38,39,41,45,47,48,52,55,58]	How clinically beneficial eHealth is perceived to be for the resident	11 (38)
Person-Centered care[b] [40,41,45,52]	Whether eHealth can be tailored to the individual needs of the resident and care home	4 (14)
Resident experience[b] [34,39,40,52,56,57]	The effect that eHealth has on resident needs and satisfaction with care	6 (21)
Outer setting: external influences on eHealth implementation		
Cosmopolitanism [37,43]	The degree to which the care home is networked with others	2 (7)
Peer pressure	Pressure experienced by the care home to implement eHealth	0
External policy and incentives [40,49,52-54]	External influences of implementation of eHealth for the care home	5 (17)
Inner setting: characteristics of the implementing care home		
Structural characteristics [35,37,43,52]	The social architecture, age, maturity, and size of the care home	4 (14)
Networks and communications [47]	The nature and quality of social networks and communication within a care home	1 (3)
Culture	Norms and values of the care home	0
Implementation climate	The capacity for change and shared receptivity of individuals to eHealth and the extent to which it will be supported within the care home	
Tension for change [32]	The extent to which stakeholders perceive current practices as needing change	1 (3)
Compatibility [36-38,40,47,51,52,54]	The degree of fit between the care home and eHealth in terms of values and existing workflows	8 (28)
Relative priority	Individuals' shared perception of the importance of implementation within the care home	0
Organizational incentives and rewards	Incentives to increase participation with eHealth such as awards and promotions for staff	0
Goals and feedback	The degree to which goals of eHealth are acted upon and feedback to staff	0
Learning climate [32,33,39,40,43,44,51,52,58,59]	A climate in which staff feel valued in the implementation process and comfortable to participate through encouragement by care home leaders	10 (34)
Readiness for implementation	Indicators of care home commitment to eHealth implementation	
Leadership engagement [31,32,37,39,43]	Commitment and involvement of care home managers and leaders in implementation	5 (17)
Available resources [32,33,35,36,39,41,42,44,46,48,51-53,59]	The level of care home resources dedicated to eHealth implementation, including money and staff time	14 (48)

CFIR[a] construct and subconstructs	Definition in the context of care homes and integrated care for people with dementia	Total, n (%)
Access to knowledge and information [31-33,35,37,38,40-44,48,51,52,54,56,59]	Access to sufficient eHealth training for end users	17 (59)
Individual characteristics: end-user individual beliefs, knowledge, and attitudes toward eHealth and implementation		
Knowledge and beliefs about the intervention [36,38,39,41,43,45-48,51-53,55,56,58,59]	End users' attitudes toward eHealth and its impact	16 (55)
Self-efficacy [33,36,38,41,47,56,59]	End users' belief in their own abilities to use the eHealth intervention	7 (24)
Individual stage of change	The phase an individual is in as they progress toward sustained use of eHealth	0
Individual identification with organization	End user's perception of their relationship with the care home	0
Other personal attributes [32,33,40,43,44,58,59]	Individuals' attributes that affect implementation such as staff willingness, experience, age, or grade	7 (24)
Process: stages of the implementation process that can impact its success		
Planning [32,38,42-44,46,48,51,54,56-59]	The degree to which tasks required for implementation of eHealth are agreed in advance	13 (45)
Engaging	Attracting and involving stakeholders in eHealth implementation	
Champions [32,35,37,43,48,58,59]	Individuals who are dedicated to driving the implementation of eHealth and overcoming resistance in the care home	7 (24)
End users[b] [32,33,37,39,40,42-44,54,59]	Other stakeholders, including end users and staff, within the care home	10 (34)
Opinion leaders	Individuals in a care home who have a formal or informal influence on others' attitudes toward implementation	0
Formally appointed internal leaders	Individuals from within the care home who are formally appointed to implement eHealth	0
External change agents	Individuals from outside the care home who formally influence implementation of eHealth	0
Executing [32,35,37,43,44]	The extent to which eHealth implementation is conducted as planned	5 (17)
Reflecting and evaluating [35,37,43,46,51,55,59]	Monitoring of eHealth implementation and feedback about its progress	7 (24)

[a]CFIR: Consolidated Framework for Implementation Research.

[b]Additional subconstructs identified inductively from the data.

Intervention Characteristics

This refers to aspects of the eHealth intervention that might affect implementation success in care homes and includes findings relating to intervention complexity, adaptability, and cost.

Relative Advantage

eHealth that is perceived as advantageous to an alternative system by improving access to emergency care [39], increasing efficiency [37,38,54,55], and reducing paperwork [47] is more likely to be adopted [48]. Barriers to uptake include a preference for face-to-face consultations [56] and increased time required to organize eHealth consultations [36].

Adaptability

eHealth is at an advantage if it aligns with data and technology already in use [37,40,51,52]. Concerns on patient privacy and electronic transfer of confidential information act as *institutional firewalls* [40,49] but can be overcome by encrypting data and assigning residents confidential numbers [49]. A way to increase the adaptability of a device is through the provision of customizable tools such as drop-down menus [40]. Incorporating a decision support system in eHealth interventions is advocated to respond to changes in individual residents' needs by providing alerts and directing staff to appropriate care [40,48,55].

Complexity

eHealth is more likely to be implemented if it is straightforward and user-friendly [36-38,40,41,44,45,47,48,53,54,56,58]. Simple devices are regarded as more reliable [36,40,48], and it is recommended that more advanced technology be used only where necessary [36]. Dual systems of paper-based and electronic devices should be avoided to minimize inconsistency [38] and the complexity of data recording [48]. Uptake is also influenced by the ease of access to eHealth [36,37,40,47,54,55], with portable tools improving ease of access [38,48] and thereby saving staff time [37,38,47,55]. However, some staff report that handheld devices are easy to break and misplace [47] and prefer a desktop computer [40]. Technological difficulties, including software and memory issues [47], inability of patients and physicians to see or hear each other [52,56], and difficulty in obtaining technical support despite frequently requiring it [36,56], can impede eHealth implementation [35,36,44,47,53,56,58]. Providing training and specialist technical support for staff are key to overcoming these barriers [40-42,44,46,53,56].

Cost

Cost is a major barrier to implementing electronic health records in care homes [52]; for an intervention to be implemented, the benefits must be perceived to outweigh the costs [36,51]. eHealth tools that incur no additional financial cost beyond staff time [37,49] and installation [54] can integrate with preassembled data [37], are cheaper to purchase than full-size computers [47], and are more likely to be adopted. Uptake is optimized by external funding [51,53,54] and intervention outcomes that minimize spending such as reduced resident transition between care settings, for example, unplanned hospitalization [36,49,53], or reduced antipsychotic prescriptions [50].

Patient Needs

This refers to the extent to which resident needs are known, prioritized, and pursued by a care home. A total of 3 novel subconstructs were identified inductively.

Clinical Benefit

An important contributor toward successful implementation is the perceived clinical benefit for residents [34,36,38,39,47,52,55,58]. eHealth interventions are reported to help staff *focus on the resident's condition* [47], better manage residents' symptoms and vital signs [34,55,58], improve safety around medication administration [55], enable staff to attend to residents in a timely manner, and identify common behavior patterns [38]. The positive impact of eHealth interventions to minimize burdensome care transitions has also been recognized. Reduced unplanned hospital transitions are suggested to improve residents' quality of life [45] by preventing disruption in care continuity with attendance to care needs in the care home [38]. Residents appreciate minimal journeys to acute care settings [52] and the benefit of emergency telemedicine sessions to deliver timely skilled health care interventions remotely [39]. Barriers to uptake of eHealth interventions pertained to opinions that the intervention little benefited residents [48] by inadequately monitoring and identifying resident change or deterioration and failing to improve assessment processes [41].

Person-Centered Care

Interventions must be tailored to the individual and accommodate changing resident needs if they are to be successfully implemented [40,41,52]. Given the heterogeneous population in care homes and variable incidence of dementia, eHealth interventions will not make a positive difference if they do not have the capacity to assist with dementia-specific needs when required [40,41,45]. Extra consideration must go into ensuring that technology is unobtrusive if used in the presence of this patient group [40,52] who may lack the capacity to understand the change in care. Where residents experience cognitive, visual, or hearing difficulties, eHealth must be tailored, such as through the provision of a larger monitor [52].

Resident Experience

A negative resident experience of eHealth obstructs implementation [39,40,57]. Dissatisfaction stems from devices interfering with the time spent with staff [57] and a switch from face-to-face to remote communication [56]. Staff report concerns that using eHealth in the presence of residents may be intrusive and dehumanizing [39,40]. Concerns tend to diminish with continued use of eHealth tools; however, over time, staff acknowledge that eHealth can improve care quality [39,52]. Other residents report no unintentional harm or negative effects on communication [57], with 1 study attributing positive findings to residents, reporting that they felt more *followed and cared for* [34].

Outer Setting

The outer setting is concerned with external influences on intervention implementation. This was least considered construct across the studies, with the main focus on the 'External policy and barriers' subconstruct. External barriers include policies on the medical liability of telemedicine [53], licensing requirements that do not allow physicians to consult across different parts of the country [49], and issues around reimbursement policies [52-54]. External financial support acts as an incentive to circumvent the additional cost barriers to implementation [40,53,54].

Inner Setting

The most commonly considered construct across the studies was the inner setting, referring to the internal characteristics of a care home. These focused on the 2 main subconstructs of implementation climate (compatibility and learning climate) and readiness for implementation (leadership, available resources, and access to knowledge).

Structural Characteristics

The structural characteristics of the care home setting that affect implementation success include the complex patient population, with the heterogeneity of residents' conditions affecting the compatibility between the intervention and the care home [35,60]. Care home size can also affect the uptake of an eHealth intervention, with larger homes facilitating uptake through more comprehensive provision of information technology services or inhibiting uptake with larger home leaders exhibiting more resistance to adoption and delivering training [43].

Implementation Climate

Compatibility

Portable eHealth devices that can be used at the point of care [38,40,54] and during nighttime hours [36] prevent disruption to workflow, thereby facilitating uptake [36,37,40,47,51,52,54]. Interventions are at an advantage if their goals are aligned with those of the care home, for example, improving advance care planning [37], and they may face resistance if they do not support existing practice [40]. Providing individually tailored implementation protocols [37] and delivering the intervention when care homes are maximally staffed can ease adoption [43,58].

Learning Climate

Creating a climate within the care home that encourages learning and active participation in the intervention is key [32,39,43,44,51,52,59]. Staff members in the frontline must be receptive [43] and participate if they are to influence patient outcomes [33,44] and affect care delivery [32]. Changes to

personnel at the site [44] and reluctance from staff [51,52,58,59] undermine a conducive learning climate. Hierarchical staffing can inhibit implementation [32,58] if junior staff feel they are unable to speak up [32] or access the new tools [40,43]. If a learning climate is not fostered, the staff lack practice in delivering an intervention [32] and opportunities to share and embed their learning to change practice [43].

Readiness for Implementation

Leadership Engagement

Organizational commitment from the top down is crucial [31,32,37,43]. Staff leaders and managers who collaborate with the research team [37] and *lead by example* through attending training and contributing to data collection [43] foster an engaged and cohesive workforce [39]. A lack of enthusiasm from management can lead to delays in implementation and reluctance from other staff [43]. Unwillingness and resistance from general practitioners to change practice also act as a barrier [32,39] and contribute to a fragmented learning climate [43].

Available Resources

Providing extra time for implementation planning [33,39,41,46,52] and end-user training [32,33,42,46,52] is required to accommodate a change in practice. A collective staff opinion of insufficient time to adjust to change hinders implementation [32,39,44,48,51,52,59]. Sufficient bandwidth is required to allow eHealth to function properly [35,36,44,53], and insufficient bandwidth can lead to significant *motion artefact* [44] and *jerky motion* images [53]. Additional space may be required for training, either within the home [42] or off site [32]. Equipment needs may include extra computers [42] and software [59], incurring further financial costs [32,51]. Reduced staffing [51], high staff turnover [44,58], and staff absence [43] obstruct the implementation of eHealth interventions. Limited resources can act as a barrier unless benefits can be demonstrated to outweigh costs [43,51]. Managers need to consider the additional resources required to accommodate adoption [43].

Access to Knowledge and Information

Providing adequate training to end users is crucial for promoting uptake [31-33,35,37,38,40-44,48,51,52,56,59]. A variety of educational methods are used, including interactive learning strategies [56], lectures, exercises, and group discussions [59]. Training is often provided through a cascaded learning *train-the-trainer* approach [37,43,48]. Preference varies as to receiving training *on the job* [31,40,42] or continuously over a designated period in the months leading up to implementation [35,37,48]. Training tailored to individual needs [38,40,48] can help facilitate uptake, whereas inadequate training can negatively influence perceptions of intervention benefits [32,48] and how staff disseminate knowledge to colleagues [32]. Incorporating a period of joint work and supervision from a qualified external expert can further facilitate the benefits of training [32].

Individual Characteristics

This pertains to end users' individual beliefs, knowledge, and attitudes toward eHealth and implementation. The findings focused on 3 main subconstructs.

Knowledge and Beliefs About the Intervention

Fostering a positive end-user attitude toward eHealth is key to ensuring willing adoption. eHealth tools that are perceived as high quality [45,48,53] and efficient [38,47,48,55], by cutting back on paperwork [47] and providing instant access to data [48], are more likely to be adopted. Interventions judged to improve staff understanding and knowledge [39,48,52,56,58] and benefit residents [36,39,48,58] are more likely to be used. eHealth devices that improve overall job satisfaction [45,48,55,56] while providing additional benefits, such as reducing staff anxiety [38], and social features that enrich staff lives outside of work [47] encourage use. Conversely, negative attitudes toward eHealth act as barriers to implementation. These include concerns regarding a lack of care improvement [39,41,45,58], a preference to deliver care in person [39,45,52,56], and concerns of inequity in health care provision, resulting in difficulty obtaining family consent [39]. Other barriers include a lack of positive impact on staff workload [36,41,45], concerns over increased time and effort because of implementation [36,39,51,52,58], and uncertainty about the purpose of eHealth and how it works [43,46]. This indicates the need for comprehensive training to highlight the benefits and importance of implementing the intervention [48].

Self-efficacy

Self-efficacy refers to an individual's belief in their own capabilities to implement an intervention. Low self-confidence in ability [47,56,59] and apprehension toward using new technology [38,47] can act as barriers to implementation. These can be overcome by providing sufficient training, which is tailored to individual needs [38,40,48], and on-hand technical support [36,56].

Other Personal Attributes

Other issues that affect implementation relate to staff willingness and commitment to ensuring that the intervention is executed as planned [32,33,44,59]. A lack of experience can also act as a barrier if senior members of staff consider more junior members to be less capable of participating [32,40,43,59] because of age, grade, educational background, or experience [32,59]. However, demonstrating that training designed for medical caregivers is effective across disciplines and grades should help dispel these notions [58].

Implementation Process

This concerned reporting on the stages of the implementation process that affect its success. A novel subconstruct was inductively identified to capture data related to engaging end users and other stakeholders.

Planning

Insufficient planning and preparation for change can result in overwhelmed staff perceiving eHealth to be incompatible with their care home [51], technology systems and facilities being

inundated [42], and essential care being omitted [46]. Poorly planned study timelines can lead to a low number of participants if enough time is not allocated to recruitment [58] and to implementation impact being undetected if insufficient time is allocated to following up and monitoring outcomes [48,57,59]. Time must also be allocated to prospectively explore different types of eHealth to ensure suitability [54]. Multiple strategies are required to initiate a change in behavior and routine practice [32,38]. Developing contingency plans to support the program [44,54] and establishing a team whose primary goal is to assist implementation can help facilitate uptake [59]. Staff members must be fully informed of the aim of the intervention [43] and have full clarity of their designated role to achieve maximum impact [56].

Engaging: Champions

Engaging individuals is key to successful implementation. This is often done through designating *champions*—individuals from within an organization responsible for driving and supporting implementation [32,35,37,43,48,58,59]. To be successful, there must be a sufficient number of assigned champions in each home [32]. They must be committed to the position [43] and not feel undermined by others in leadership roles [32]. Designating champions is often undertaken by the team manager [59] to ensure they are appropriately able to support other staff members [43,48].

Engaging: End Users

Engaging champions alone is insufficient; other staff members must be involved for successful implementation [33,37,39,40,42-44,54,59]. For a change in the quality of care to be observed, staff engagement must be sustained [43], with all disciplines, including frontline staff [33,54] and corporate leaders, committed to implementation [37]. Involving end users in the design and development of technology [39,40,59] and considering individual requirements is key if an intervention is to be embedded in a new context.

Execution

Low intervention fidelity is a challenge in complex care home settings [35], and residents often do not receive the intervention as planned [32]. Flexible protocols tailored to the care home environment should be developed [37,44]. Barriers to successful execution include lengthy gaps between initial contact with participants and data collection and intervention activity, which can undermine the coherence of the intervention [43].

Reflecting and Evaluating

Ongoing evaluation of progress is vital to ensure a new intervention is effectively embedded [37,43,46,51,55,59]. This can be done using routine data collected through the eHealth tool [43], through in-person visits to the home, monthly conference calls, or video status reports [35,37] to determine the effect and use of the intervention [46]. Receiving feedback can also motivate care home participation, with an absence of evaluative feedback leading to a feeling of being *short-changed* [43].

Discussion

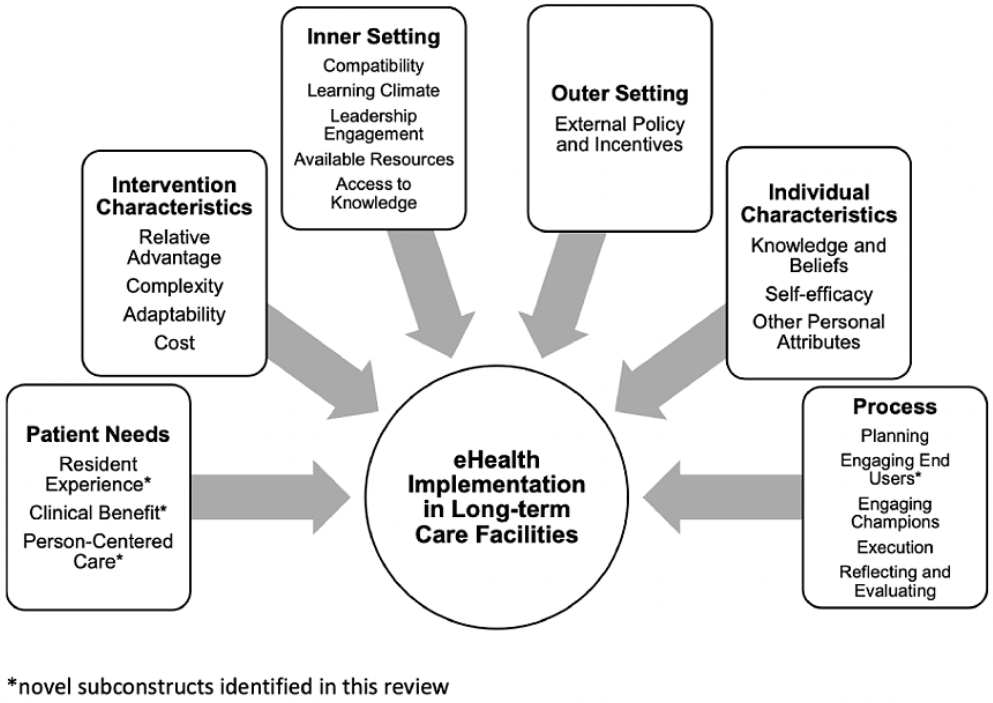
Principal Findings

This review investigated the key facilitators and barriers that influence the implementation of eHealth interventions to support comprehensive assessment and decision-making for people with dementia in care homes. Our findings inform a model of eHealth implementation in long-term care for people with dementia (Figure 2). We identified 4 novel subconstructs (denoted in the figure by an asterisk) required to adapt the CFIR to the care home context and use of eHealth to enhance integrated care for people with dementia. This modification enables the framework to accommodate the context-specific findings identified in this review and enhances its use when applied to implement eHealth in the care home context.

The 3 novel subconstructs—*resident experience*, *clinical benefit*, and *person-centered care*—were identified within the *patient needs* construct. No subconstructs have previously been delineated here, and the framework was previously not nuanced enough to capture the data and implementation requirements for this population. *Clinical benefit* was the most commonly identified theme, with the capacity to minimize burdensome transition between care settings as a key facilitator of implementation. This is echoed in the literature that highlights the adverse impact of burdensome care transitions, both on the resident [61] and on costs [62]. Although outcomes were not the focus of this review, eHealth interventions are more likely to be embedded if they can improve the delivery of integrated health care where it is required, enhance integrated care, and improve cost-effectiveness in the National Health Service [63].

Findings around the potential dehumanization of care reiterate a concern raised previously in relation to eHealth implementation [64]. Although the aim of eHealth is to enhance care delivery, entirely omitting in-person contact has clear disadvantages. Striking the right balance between delivering care face to face and virtually is crucial. Future research should focus on delineating the components of health care that must be delivered in person and identifying those that are more amenable to remote delivery to ensure that care quality is not compromised.

Figure 2. A model of factors that influence the implementation of eHealth in care for people with dementia.



The *engaging* subconstruct of the *implementation process* was also modified for the purpose of this review. None of the proposed subcategories of *engaging* (opinion leaders, formally appointed internal implementation leaders, champions, and external change leaders) were suitable for capturing data around engaging end users. The importance of involving end users throughout the implementation process is a key finding and is consistently advocated in the literature to help close the gap between research production and clinical practice [65,66]. Therefore, a novel subcategory titled *end users* was added to the CFIR. Contradictory findings regarding staff and resident preferences on the design and use of eHealth devices are an important indication that there is no *one-size-fits-all* approach to eHealth implementation. This inconsistency in preferences highlights the critical need to involve end users during intervention development to accommodate a range of requirements and enhance the likelihood of sustained implementation. Progressing the CFIR to include a subcategory suitable for data pertaining to all stakeholders' input increases its relevance and applicability for this context.

The most salient construct identified in this review was *inner setting* of the implementing organization. This is concurrent with the existing research on care homes. A systematic review by Goodman et al [67] identified many of the same factors that influence the readiness of care homes to participate in change. These include ensuring compatibility between the intervention and existing care home routine, providing sufficient training and resources, and engaging care home leaders. Both reviews highlight the significance of the organizational context on implementation success and the need to consider context in the planning stage to inform study design rather than retrospectively. This is contrary to a review of determinants for eHealth implementation with informal caregivers in the community [68], where few factors relating to context were identified as important. This indicates that factors influencing implementation may not be consistent across health care settings and stresses the importance of understanding specific contextual factors and tailoring implementation strategies accordingly.

One of the most prominent determinants of implementation identified in this review was the complexity of the intervention, consistent with previous findings highlighting ease of use as the most important facilitator of eHealth adoption. The nonadoption, abandonment, scale-up, spread, and sustainability framework [69] was developed specifically in response to the finding that eHealth, which is categorized as complex is rarely, if ever, successfully embedded in mainstream care. It aims to help innovators and implementers measure and minimize complexity in eHealth and scale-up and sustain innovation, and therefore, it could helpfully be used in this context.

Implications for Policy and Clinical Practice

Using the most salient subconstructs of the framework, a conceptual model was developed to highlight the most important factors that influence the implementation of eHealth interventions to enhance integrated care focusing on care processes of comprehensive assessment and decision-making about care and treatment (Figure 2). The findings can be translated into practical recommendations for organizations aiming to embed eHealth within long-term care settings for people with dementia (Textbox 2). The model indicates that for implementation to be successful, eHealth devices must be low cost, simple to use, and tailored to the care home setting and residents. It must be clinically beneficial to the residents, with special consideration of changing, multimorbid dementia-specific needs. Extensive planning and engagement of care home leaders, end users, and champions in the

development and implementation process are key to ensuring successful execution. Providing sufficient training and resources to ensure that care home staff feel valued, motivated, and optimistic about a change in practice is crucial to fostering a positive implementation climate.

Textbox 2. Practical recommendation for implementation of eHealth in long-term care for residents with dementia.

Consolidated Framework for Implementation Research constructs and practical recommendations for eHealth implementation

Patient needs

- eHealth should be tailored to the individual resident and accommodate changing and complex needs.
- It should be unobtrusive and not replace in-person contact or compromise care quality.
- eHealth must not just streamline workload but clinically benefit the resident and improve outcomes.

Intervention characteristics

- eHealth should be user-friendly and accessible to increase sustainability.
- eHealth that is interoperable with current systems is advisable to minimize installation cost and complex new learning for end users.
- Technical support should be easily attainable and readily available.

Outer setting

- Policies that endorse eHealth use in care homes ease implementation, for example, rapid policy-driven implementation of remote consultations during the COVID-19 pandemic.

Inner setting

- eHealth tools should be tailored to care home settings to fit with existing workflow and care home values.
- Engaging care home leaders in the intervention is key to promoting enthusiasm and a cohesive working environment.
- Staff participation and learning about eHealth should be encouraged.
- Sufficient training should be provided for end users, which should be tailored to individual requirements.
- Additional resources need to be allocated to accommodate a change in practice. These include technological requirements such as bandwidth and equipment and extra staff time.

Individual characteristics

- Fostering a positive staff attitude toward eHealth is essential for uptake.
- The tool should benefit staff by easing workload, increasing knowledge, and improving job satisfaction.

Implementation process

- Preparation for change and consideration of implementation timelines, strategies, and contingency plans in advance is crucial.
- *Champions* should be designated in each care home to drive and support implementation.
- End users and other stakeholders should be engaged from an early point and consulted in the implementation process to accommodate individual requirements.
- Ongoing evaluation and reflection on uptake and adherence to eHealth should occur to inform any necessary developments and improvements.

CFIR Framework

The CFIR is a comprehensive determinant framework and is chosen to guide analysis owing to its previous application with eHealth interventions [68,70,71]. Generally, the extracted data were amenable to the CFIR, with no data left uncoded. However, 33% (13/39) of the subconstructs had no associated data (Table 2), concurrent with findings from a previous review of eHealth implementation across health care settings [70]. This consistency either suggests that some subconstructs are not relevant to implementing eHealth or highlights a limitation of the evidence base and lack of existing literature for this setting and population. These areas of uncertainty are important for informing future research, which should focus on identifying barriers to and facilitators of the implementation of these underresearched constructs.

Strengths and Limitations

A systematic approach to this review allowed for the rigorous and thorough identification of the relevant literature. Evidence synthesis was theory driven and guided by the CFIR, which has been built upon here to increase its relevance in eHealth and long-term care contexts.

Although this review specifies care home residents with dementia as its population of interest, only 38% (11/29) of the included studies had an exclusive dementia population. This was to reflect the real-world heterogeneous populations in care homes and include interventions that were suitable at an

organizational level rather than the individual level. Although there was no real difference in factors that influence implementation between dementia-specific and mixed populations, caution must be exercised when extrapolating these findings to homogeneous dementia populations.

Gray literature was not included in this review. This was because, when identified, it provided little data on the factors that influence implementation. This could have excluded relevant data on the eHealth interventions used in care homes. A recent scoping review on telehealth during the COVID-19 pandemic reported a rapid rise in eHealth during the pandemic [72].

Conclusions

To our knowledge, this is the first review to synthesize evidence on the implementation of eHealth interventions focusing specifically on improving assessment and decision-making in care homes for people with dementia. We adapted the CFIR and progressed its applicability for use in this context. We developed a conceptual model to demonstrate the most important factors to consider when designing and implementing an eHealth intervention and translated it into 18 practical recommendations for implementers, innovators, and organizations to implement eHealth for people with dementia in long-term care. Particular focus should be placed on the individual care home setting and on the consideration of resident and end-user needs when developing an implementation strategy for use in this context.

Acknowledgments

The authors would like to thank Adam Garside, an information support specialist at King's College London, who helped with the development of the search strategy. This project was funded by the Economic and Social Research Council (ESRC), National Institute for Health Research (NIHR), and ESRC/NIHR Dementia Initiative 2018 (grant reference number ES/S010327/1). CE is funded by a Health Education England/NIHR Senior Clinical Lectureship (ICA-SCL-2015-01-001). ND was supported by a fellowship from the Alzheimer's Society, United Kingdom (grant number AS-JF-16b-012). The views expressed are those of the authors and not necessarily those of the ESRC, NIHR, or the Department of Health and Social Care.

Authors' Contributions

JG, CE, ND, and JEA designed the study. JG, EY, and JA contributed to data screening, extraction, and analysis. JG wrote the manuscript. The authors contributed to revisions and read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[PDF File (Adobe PDF File), 96 KB - jmir_v24i2e29837_app1.pdf]

Multimedia Appendix 2
Search strategy.
[DOCX File , 18 KB - jmir_v24i2e29837_app2.docx]

Multimedia Appendix 3
Quality appraisal summary table.
[DOCX File , 16 KB - jmir_v24i2e29837_app3.docx]

Multimedia Appendix 4
Differences in implementation requirements.
[DOCX File , 18 KB - jmir_v24i2e29837_app4.docx]

References

1. Deaths registered in England and Wales: 2019. Office of National Statistics. 2020. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsregistrationsummarytables/2019 [accessed 2021-12-10]
2. Sleeman KE, de Brito M, Etkind S, Nkhoma K, Guo P, Higginson IJ, et al. The escalating global burden of serious health-related suffering: projections to 2060 by world regions, age groups, and health conditions. Lancet Glob Health 2019 Jul;7(7):e883-e892 [FREE Full text] [doi: 10.1016/S2214-109X(19)30172-X] [Medline: 31129125]
3. Sampson EL, Burns A, Richards M. Improving end-of-life care for people with dementia. Br J Psychiatry 2011 Nov;199(5):357-359. [doi: 10.1192/bjp.bp.111.097030] [Medline: 22045942]

4. Tonelli M, Wiebe N, Straus S, Fortin M, Guthrie B, James MT, Alberta Kidney Disease Network. Multimorbidity, dementia and health care in older people: a population-based cohort study. CMAJ Open 2017 Aug 14;5(3):E623-E631 [FREE Full text] [doi: 10.9778/cmajo.20170052] [Medline: 28811281]
5. Leniz J, Higginson I, Stewart R, Sleeman K. Understanding which people with dementia are at risk of inappropriate care and avoidable transitions to hospital near the end-of-life: a retrospective cohort study. Age Ageing 2019 Sep 01;48(5):672-679. [doi: 10.1093/ageing/afz052] [Medline: 31135024]
6. Sampson EL, Candy B, Davis S, Gola AB, Harrington J, King M, et al. Living and dying with advanced dementia: a prospective cohort study of symptoms, service use and care at the end of life. Palliat Med 2018 Mar;32(3):668-681 [FREE Full text] [doi: 10.1177/0269216317726443] [Medline: 28922625]
7. Hendriks SA, Smalbrugge M, Hertogh CM, van der Steen JT. Dying with dementia: symptoms, treatment, and quality of life in the last week of life. J Pain Symptom Manage 2014 Apr;47(4):710-720 [FREE Full text] [doi: 10.1016/j.jpainsymman.2013.05.015] [Medline: 23916680]
8. Data Analysis Report Dying with Dementia. Public Health England. 2016. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/733065/Dying_with_dementia_data_analysis_report.pdf [accessed 2021-12-11]
9. End of life care in frailty: care homes. British Geriatrics Society. 2020. URL: https://www.bgs.org.uk/resources/end-of-life-care-in-frailty-care-homes [accessed 2021-12-11]
10. Understaffed and overworked: poor working conditions and quality of care in residential care facilities for the elderly. Publications. 2017. URL: https://digitalcommons.law.ggu.edu/pubs/788/ [accessed 2021-12-11]
11. Briggs AM, Valentijn PP, Thiyagarajan JA, Araujo de Carvalho I. Elements of integrated care approaches for older people: a review of reviews. BMJ Open 2018 Apr 07;8(4):e021194 [FREE Full text] [doi: 10.1136/bmjopen-2017-021194] [Medline: 29627819]
12. Ellis G, Whitehead MA, O'Neill D, Langhorne P, Robinson D. Comprehensive geriatric assessment for older adults admitted to hospital. Cochrane Database Syst Rev 2011 Jul 06(7):CD006211 [FREE Full text] [doi: 10.1002/14651858.CD006211.pub2] [Medline: 21735403]
13. Disalvo D, Luckett T, Bennett A, Davidson PM, Agar M. Multidisciplinary perspectives on medication-related decision-making for people with advanced dementia living in long-term care: a critical incident analysis. Eur J Clin Pharmacol 2020 Apr;76(4):567-578. [doi: 10.1007/s00228-019-02820-z] [Medline: 31932874]
14. Coombs MA, Parker R, de Vries K. Managing risk during care transitions when approaching end of life: a qualitative study of patients' and health care professionals' decision making. Palliat Med 2017 Jul;31(7):617-624 [FREE Full text] [doi: 10.1177/0269216316673476] [Medline: 28618896]
15. Monaghesh E, Hajizadeh A. The role of telehealth during COVID-19 outbreak: a systematic review based on current evidence. BMC Public Health 2020 Aug 01;20(1):1193 [FREE Full text] [doi: 10.1186/s12889-020-09301-4] [Medline: 32738884]
16. Eysenbach G. What is e-health? J Med Internet Res 2001 Jun 18;3(2):E20 [FREE Full text] [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
17. Coronavirus (COVID-19): notification to organisations to share information. GOV.UK. URL: https://www.gov.uk/government/publications/coronavirus-covid-19-notification-of-data-controllers-to-share-information [accessed 2021-12-11]
18. Hanratty B, Burton JK, Goodman C, Gordon AL, Spilsbury K. Covid-19 and lack of linked datasets for care homes. BMJ 2020 Jun 24;369:m2463. [doi: 10.1136/bmj.m2463] [Medline: 32581008]
19. Comas-Herrera A, Ashcroft E, Lorenz-Dant K. International examples of measures to prevent and manage COVID-19 outbreaks in residential care and nursing home settings. International Long Term Care Policy Network. URL: https://ltccovid.org/2020/05/02/new-report-international-examples-of-measures-to-prevent-and-manage-covid-19-outbreaks-in-residential-care-and-nursing-home-settings/ [accessed 2021-12-11]
20. Haralambous B, Subramaniam S, Hwang K, Dow B, LoGiudice D. A narrative review of the evidence regarding the use of telemedicine to deliver video-interpreting during dementia assessments for older people. Asia Pac Psychiatry 2019 Sep;11(3):e12355. [doi: 10.1111/appy.12355] [Medline: 31025506]
21. Schaller S, Marinova-Schmidt V, Setzer M, Kondylakis H, Griebel L, Sedlmayr M, et al. Usefulness of a tailored eHealth service for informal caregivers and professionals in the dementia treatment and care setting: the eHealthMonitor dementia portal. JMIR Res Protoc 2016 Apr 05;5(2):e47 [FREE Full text] [doi: 10.2196/resprot.4354] [Medline: 27050401]
22. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M. Guidance on the Conduct of Narrative Synthesis in Systematic Reviews: A Product From the ESRC Methods Programme. United Kingdom: Lancaster University; 2006.
23. Methley AM, Campbell S, Chew-Graham C, McNally R, Cheraghi-Sohi S. PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. BMC Health Serv Res 2014 Nov 21;14:579 [FREE Full text] [doi: 10.1186/s12913-014-0579-0] [Medline: 25413154]
24. Hupe M. EndNote X9. J Electronic Resources Med Libraries 2019 Nov 26;16(3-4):117-119. [doi: 10.1080/15424065.2019.1691963]
25. CASP Checklists. CASP-UK. URL: https://casp-uk.net/casp-tools-checklists/ [accessed 2021-12-11]

26. Hong QN, Fàbregues S, Bartlett G, Boardman F, Cargo M, Dagenais P, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. Educ Inf 2018 Dec 18;34(4):285-291. [doi: 10.3233/efi-180221]
27. Critical appraisal tools. University of Adelaide. URL: https://joannabriggs.org/critical-appraisal-tools [accessed 2021-12-11]
28. Safaeinili N, Brown-Johnson C, Shaw JG, Mahoney M, Winget M. CFIR simplified: pragmatic application of and adaptations to the Consolidated Framework for Implementation Research (CFIR) for evaluation of a patient-centered care transformation within a learning health system. Learn Health Syst 2020 Sep 26;4(1):e10201 [FREE Full text] [doi: 10.1002/lrh2.10201] [Medline: 31989028]
29. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009 Aug 07;4:50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
30. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
31. Mitchell SL, Shaffer ML, Cohen S, Hanson LC, Habtemariam D, Volandes AE. An advance care planning video decision support tool for nursing home residents with advanced dementia: a cluster randomized clinical trial. JAMA Intern Med 2018 Jul 01;178(7):961-969 [FREE Full text] [doi: 10.1001/jamainternmed.2018.1506] [Medline: 29868778]
32. Moniz-Cook E, Hart C, Woods B, Whitaker C, James I, Russell I. Challenge Demcare: management of challenging behaviour in dementia at home and in care homes. Programme Grants Appl Res 2017;5(15). [doi: 10.3310/pgfar05150]
33. Daly JM, Buckwalter K, Maas M. Written and computerized care plans. Organizational processes and effect on patient outcomes. J Gerontol Nurs 2002 Sep;28(9):14-23. [doi: 10.3928/0098-9134-20020901-05] [Medline: 12240516]
34. De Luca R, Bramanti A, De Cola MC, Trifiletti A, Tomasello P, Torrisi M, et al. Tele-health-care in the elderly living in nursing home: the first Sicilian multimodal approach. Aging Clin Exp Res 2016 Aug;28(4):753-759. [doi: 10.1007/s40520-015-0463-8] [Medline: 26420423]
35. Mitchell SL, Volandes AE, Gutman R, Gozalo PL, Ogarek JA, Loomer L, et al. Advance care planning video intervention among long-stay nursing home residents: a pragmatic cluster randomized clinical trial. JAMA Intern Med 2020 Aug 01;180(8):1070-1078 [FREE Full text] [doi: 10.1001/jamainternmed.2020.2366] [Medline: 32628258]
36. Weiner M, Schadow G, Lindbergh D, Warvel J, Abernathy G, Perkins S, et al. Clinicians' and patients' experiences and satisfaction with unscheduled, nighttime, internet-based video conferencing for assessing acute medical problems in a nursing facility. AMIA Annu Symp Proc 2003:709-713 [FREE Full text] [Medline: 14728265]
37. Mor V, Volandes AE, Gutman R, Gatsonis C, Mitchell SL. PRagmatic trial Of Video Education in Nursing homes: the design and rationale for a pragmatic cluster randomized trial in the nursing home setting. Clin Trials 2017 Apr;14(2):140-151 [FREE Full text] [doi: 10.1177/1740774516685298] [Medline: 28068789]
38. Klein B, Clinnick L, Chesler J, Stranieri A, Bignold A, Dazeley R, et al. Supporting regional aged care nursing staff to manage residents' behavioural and psychological symptoms of dementia, in real time, using the Nurses' Behavioural Assistant (NBA): a pilot site 'end-user attitudes' trial. Stud Health Technol Inform 2018;246:24-28. [Medline: 29507257]
39. Piau A, Vautier C, De Mauleon A, Tchalla A, Rumeau P, Nourhashemi F, et al. Health workers perception on telemedicine in management of neuropsychiatric symptoms in long-term care facilities: two years follow-up. Geriatr Nurs 2020;41(6):1000-1005 [FREE Full text] [doi: 10.1016/j.gerinurse.2020.07.009] [Medline: 32771311]
40. Shiells K, Diaz Baquero AA, Štěpánková O, Holmerová I. Staff perspectives on the usability of electronic patient records for planning and delivering dementia care in nursing homes: a multiple case study. BMC Med Inform Decis Mak 2020 Jul 13;20(1):159 [FREE Full text] [doi: 10.1186/s12911-020-01160-8] [Medline: 32660474]
41. Vuorinen M. Registered Nurses' experiences with, and feelings and attitudes towards, the International Resident Assessment Instrument for long-term care facilities in New Zealand in 2017. J Res Nurs 2020 Mar;25(2):141-155 [FREE Full text] [doi: 10.1177/1744987119890651 26] [Medline: 34394619]
42. Alexander G, Rantz M, Flesner M, Diekemper M, Siem C. Clinical information systems in nursing homes: an evaluation of initial implementation strategies. Comput Inform Nurs 2007;25(4):189-197. [doi: 10.1097/01.NCN.0000280589.28067.18] [Medline: 17625399]
43. Keenan J, Poland F, Manthorpe J, Hart C, Moniz-Cook E. Implementing e-learning and e-tools for care home staff supporting residents with dementia and challenging behaviour: a process evaluation of the ResCare study using normalisation process theory. Dementia (London) 2020 Jul;19(5):1604-1620 [FREE Full text] [doi: 10.1177/1471301218803195] [Medline: 30269534]
44. Johnston D, Jones BN. Telepsychiatry consultations to a rural nursing facility: a 2-year experience. J Geriatr Psychiatry Neurol 2001;14(2):72-75. [doi: 10.1177/089198870101400205] [Medline: 11419570]
45. Salles N, Lafargue A, Cressot V, Glenisson L, Barateau M, Thiel E, et al. Global geriatric evaluation is feasible during interactive telemedicine in nursing homes. European Res Telemedicine / La Recherche Européenne en Télémédecine 2017 Jul;6(2):59-65. [doi: 10.1016/j.eurtel.2017.06.002]
46. Alexander G. A descriptive analysis of a nursing home clinical information system with decision support. Perspect Health Inf Manag 2008 Sep 02;5:12 [FREE Full text] [Medline: 18784850]

47. Qadri S, Wang J, Ruiz J, Roos B. Personal digital assistants as point-of-care tools in long-term care facilities: a pilot study. Educ Gerontol 2009 Mar 10;35(4):296-307. [doi: 10.1080/03601270802467049]
48. Munyisia EN, Yu P, Hailey D. The changes in caregivers' perceptions about the quality of information and benefits of nursing documentation associated with the introduction of an electronic documentation system in a nursing home. Int J Med Inform 2011 Feb;80(2):116-126. [doi: 10.1016/j.ijmedinf.2010.10.011] [Medline: 21242104]
49. Catic AG, Mattison ML, Bakaev I, Morgan M, Monti SM, Lipsitz L. ECHO-AGE: an innovative model of geriatric care for long-term care residents with dementia and behavioral issues. J Am Med Dir Assoc 2014 Dec;15(12):938-942. [doi: 10.1016/j.jamda.2014.08.014] [Medline: 25306294]
50. Gordon SE, Dufour AB, Monti SM, Mattison ML, Catic AG, Thomas CP, et al. Impact of a videoconference educational intervention on physical restraint and antipsychotic use in nursing homes: results from the ECHO-AGE pilot study. J Am Med Dir Assoc 2016 Jun 01;17(6):553-556 [FREE Full text] [doi: 10.1016/j.jamda.2016.03.002] [Medline: 27161317]
51. Bjarnadottir R, Herzig C, Travers J, Castle N, Stone P. Implementation of electronic health records in US nursing homes. Comput Inform Nurs 2017 Aug;35(8):417-424 [FREE Full text] [doi: 10.1097/CIN.0000000000000344] [Medline: 28800581]
52. Wakefield B, Buresh K, Flanagan J, Kienzle M. Interactive video specialty consultations in long-term care. J Am Geriatr Soc 2004 May;52(5):789-793. [doi: 10.1111/j.1532-5415.2004.52220.x] [Medline: 15086663]
53. Lee JH, Kim JH, Jhoo JH, Lee KU, Kim KW, Lee DY, et al. A telemedicine system as a care modality for dementia patients in Korea. Alzheimer Dis Assoc Disord 2000;14(2):94-101. [doi: 10.1097/00002093-200004000-00007] [Medline: 10850748]
54. Lyketsos CG, Roques C, Hovanec L, Jones BN. Telemedicine use and the reduction of psychiatric admissions from a long-term care facility. J Geriatr Psychiatry Neurol 2001;14(2):76-79. [doi: 10.1177/089198870101400206] [Medline: 11419571]
55. Krüger K, Strand L, Geitung J, Eide GE, Grimsmo A. Can electronic tools help improve nursing home quality? ISRN Nurs 2011;2011:208142 [FREE Full text] [doi: 10.5402/2011/208142] [Medline: 22013540]
56. Perri G, Abdel-Malek N, Bandali A, Grosbein H, Gardner S. Early integration of palliative care in a long-term care home: a telemedicine feasibility pilot study. Pall Supp Care 2020 Feb 18;18(4):460-467. [doi: 10.1017/s1478951520000012]
57. Pillemer K, Meador RH, Teresi JA, Chen EK, Henderson CR, Lachs MS, et al. Effects of electronic health information technology implementation on nursing home resident outcomes. J Aging Health 2012 Feb;24(1):92-112 [FREE Full text] [doi: 10.1177/0898264311408899] [Medline: 21646551]
58. O'Mahony S, Marthone P, Santoro G, Horn C, Selikson S, Powell T, et al. Preliminary report of a pilot tele-health palliative care and bioethics program for residents in skilled nursing facilities. Palliat Care 2009 Dec 08;3:PCRT.S3296. [doi: 10.4137/pcrt.s3296]
59. Fossum M, Alexander GL, Ehnfors M, Ehrenberg A. Effects of a computerized decision support system on pressure ulcers and malnutrition in nursing homes for the elderly. Int J Med Inform 2011 Sep;80(9):607-617. [doi: 10.1016/j.ijmedinf.2011.06.009] [Medline: 21783409]
60. Shiells K, Holmerova I, Steffl M, Stepankova O. Electronic patient records as a tool to facilitate care provision in nursing homes: an integrative review. Inform Health Soc Care 2019 Sep;44(3):262-277. [doi: 10.1080/17538157.2018.1496091] [Medline: 30125140]
61. Dwyer R, Gabbe B, Stoelwinder JU, Lowthian J. A systematic review of outcomes following emergency transfer to hospital for residents of aged care facilities. Age Ageing 2014 Nov;43(6):759-766. [doi: 10.1093/ageing/afu117] [Medline: 25315230]
62. Wolters A, Santos F, Lloyd T, Lilburne C, Steventon A. Emergency admissions to hospital from care homes: how often and what for? The Health Foundation. 2019. URL: https://www.health.org.uk/publications/reports/emergency-admissions-to-hospital-from-care-homes [accessed 2021-12-11]
63. Delivering the benefits of digital health care. Nuffield Trust. 2016. URL: https://www.nuffieldtrust.org.uk/files/2017-01/delivering-the-benefits-of-digital-technology-web-final.pdf [accessed 2021-12-11]
64. Bailey J. Does health information technology dehumanize health care? Virtual Mentor 2011 Mar 01;13(3):181-185 [FREE Full text] [doi: 10.1001/virtualmentor.2011.13.3.msoc1-1103] [Medline: 23127322]
65. Versluis A, van Luenen S, Meijer E, Honkoop PJ, Pinnock H, Mohr DC, et al. SERIES: eHealth in primary care. Part 4: addressing the challenges of implementation. Eur J Gen Pract 2020 Dec;26(1):140-145 [FREE Full text] [doi: 10.1080/13814788.2020.1826431] [Medline: 33025820]
66. Slattery P, Saeri AK, Bragge P. Research co-design in health: a rapid overview of reviews. Health Res Policy Syst 2020 Feb 11;18(1):17 [FREE Full text] [doi: 10.1186/s12961-020-0528-9] [Medline: 32046728]
67. Goodman C, Sharpe R, Russell C, Meyer J, Gordon A, Dening T. Care home readiness: a rapid review and consensus workshops on how organisational context affects care home engagement with health care innovation. University of Hertfordshire Research Archive. 2017. URL: https://uhra.herts.ac.uk/handle/2299/18200 [accessed 2021-12-11]
68. Christie HL, Bartels SL, Boots LM, Tange HJ, Verhey FR, de Vugt ME. A systematic review on the implementation of eHealth interventions for informal caregivers of people with dementia. Internet Interv 2018 Sep;13:51-59 [FREE Full text] [doi: 10.1016/j.invent.2018.07.002] [Medline: 30206519]
69. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]

70. Ross J, Stevenson F, Lau R, Murray E. Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). Implement Sci 2016 Oct 26;11(1):146 [FREE Full text] [doi: 10.1186/s13012-016-0510-7] [Medline: 27782832]
71. Ware P, Ross HJ, Cafazzo JA, Laporte A, Gordon K, Seto E. Evaluating the implementation of a mobile phone-based telemonitoring program: longitudinal study guided by the consolidated framework for implementation research. JMIR Mhealth Uhealth 2018 Jul 31;6(7):e10768 [FREE Full text] [doi: 10.2196/10768] [Medline: 30064970]
72. Doraiswamy S, Abraham A, Mamtani R, Cheema S. Use of telehealth during the COVID-19 pandemic: scoping review. J Med Internet Res 2020 Dec 01;22(12):e24087 [FREE Full text] [doi: 10.2196/24087] [Medline: 33147166]

Abbreviations

CASP: Critical Appraisal Skills Program
CFIR: Consolidated Framework for Implementation Research
ESRC: Economic and Social Research Council
JBI: Joanna Briggs Institute
MMAT: Mixed Methods Appraisal Tool
NIHR: National Institute for Health Research
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-analyses
PROSPERO: International Prospective Register of Systematic Reviews

Edited by R Kukafka; submitted 23.04.21; peer-reviewed by D Valdes, S Lundell; comments to author 20.07.21; revised version received 29.10.21; accepted 02.11.21; published 03.02.22.

Please cite as:
Gillam J, Davies N, Aworinde J, Yorganci E, Anderson JE, Evans C
Implementation of eHealth to Support Assessment and Decision-making for Residents With Dementia in Long-term Care: Systematic Review
J Med Internet Res 2022;24(2):e29837
URL: https://www.jmir.org/2022/2/e29837
doi:10.2196/29837
PMID:35113029

Review

Attrition Within Digital Health Interventions for People With Multiple Sclerosis: Systematic Review and Meta-analysis

William Bevens[1], BSc, MSc; Tracey Weiland[1], BBSc, PhD; Kathleen Gray[2], PhD; George Jelinek[1], MD; Sandra Neate[1], MBBS; Steve Simpson-Yap[1], MPH, PhD

[1]Centre for Epidemiology and Biostatistics, The University of Melbourne, Carlton, Australia

[2]Centre for Digital Transformation of Health, The University of Melbourne, Carlton, Australia

Corresponding Author:
William Bevens, BSc, MSc
Centre for Epidemiology and Biostatistics
The University of Melbourne
207 Bouverie Street
Carlton, 3053
Australia
Phone: 61 0498337231
Email: william.bevens@unimelb.edu.au

Abstract

Background: Digital health interventions have revolutionized multiple sclerosis (MS) care by supporting people with MS to better self-manage their disease. It is now understood that the technological elements that comprise this category of digital health interventions can influence participant engagement in self-management programs, and people with MS can experience significant barriers, influenced by these elements, to remaining engaged during a period of learning. It is essential to explore the influence of technological elements in mitigating attrition.

Objective: This study aimed to examine the study design and technological elements of documented digital health interventions targeted at people with MS—digital health interventions that were intended to support a program of engagement over a defined period—and to explore how these correlated with attrition among participants of randomized controlled trials (RCTs).

Methods: We conducted a systematic review and meta-analysis of RCTs (n=32) describing digital health self-management interventions for people with MS. We analyzed attrition in included studies, using a random-effects model and meta-regression to measure the association between potential moderators.

Results: There were no measured differences in attrition between the intervention and control arms; however, some of the heterogeneity observed was explained by the composite technological element score. The pooled attrition rates for the intervention and control arms were 14.7% and 15.6%, respectively.

Conclusions: This paper provides insight into the technological composition of digital health interventions designed for people with MS and describes the degree of attrition in both study arms. This paper will aid in the design of future studies in this area, particularly for digital health interventions of this type.

(J Med Internet Res 2022;24(2):e27735) doi:10.2196/27735

KEYWORDS

digital health; meta-analysis; self-management; eHealth; attrition; digital health interventions; DHI; multiple sclerosis; MS; randomized controlled trials

Introduction

Multiple sclerosis (MS) is a chronic neurological disease that affects over 2 million people globally [1]. The disease course of MS is highly variable and can be associated with a progressive decline in physical and cognitive function. The current treatment for MS involves the use of disease-modifying treatments and symptom management; however, the delivery of health care for MS is becoming increasingly supported by digital health interventions.

People with MS readily seek out web-based information on their disease and appear to do so at rates higher than those with other neurological conditions [2,3]. Further, many people with MS engage with digital technologies to manage their disease,

including exchanging medical information with health care providers or to assist in making and maintaining positive lifestyle changes [4]. Digital health interventions that support learning to self-manage MS and scaffold learning and practicing over a specific period of time are a particular subset of digital health interventions for people with MS. Evidence suggests that there are barriers to remaining engaged for the recommended duration in these interventions [5] and technological elements that could support engagement (such as interactivity, multimedia components, or feedback) can play a pivotal role in ameliorating attrition. Understanding the extent to which these technological elements may impact engagement by people with MS in such digital health interventions remains to be explored and is a necessary next step in designing and evaluating such digital health interventions.

We conducted a meta-analysis to address the following primary questions: (1) is there a difference in attrition for participants that were allocated to the intervention in randomized controlled trials (RCTs) of digital health interventions for people with MS between intervention and control arms? (2) How do study characteristics and technological elements of digital health interventions influence observed degrees of attrition?

Methods

Eligibility Criteria

Studies were eligible for inclusion in the review if they met each of the following predetermined criteria: (1) describe RCTs examining a treatment intervention or interventions for MS; (2) were delivered via a technological platform (ie, a "computer-assisted" or "web/internet-based intervention" or "eHealth/mHealth"), limited to a PC or to Mac, smartphone, and tablet devices. Interactivity and communication over a specified period of time were important components; therefore, episodic teleconferencing or telemonitoring and virtual reality–only interventions were excluded; (3) included study participants with a diagnosis of MS (any type); (4) reported attrition for those allocated to the intervention or control group upon conclusion of the intervention and study period, and the intervention and control arms were randomized from the same subject population; that is, no comparisons with healthy controls.

Search Strategy

A literature search was conducted on April 30, 2021, for published studies in the following databases: IEEE, Medline, Scopus, and CINAHL. The search contained terms related to identifying online interventions: (*online* OR *web-based* OR *internet* OR *digital* OR *virtual* OR *computer-assisted* OR *mhealth* OR *mobile* OR *smartphone* OR *ehealth* OR *telehealth* OR *telemedicine* OR *app*), a term identifying multiple sclerosis (*multiple sclerosis*), and a Medical Subject Headings (MeSH) term (*telemedicine*).

Reference lists and in-text citations were also screened for additional studies, as well as recommendations via correspondence from authors of other included studies.

Selection and Data Collection

Duplicates were first removed within and between databases. The remaining articles were then screened on the basis of their titles and abstract; thereafter, the eligibility of the final papers was confirmed following review of the full-text articles (Figure 1). Extraction of study characteristics and outcome data was conducted by authors WB and TW.

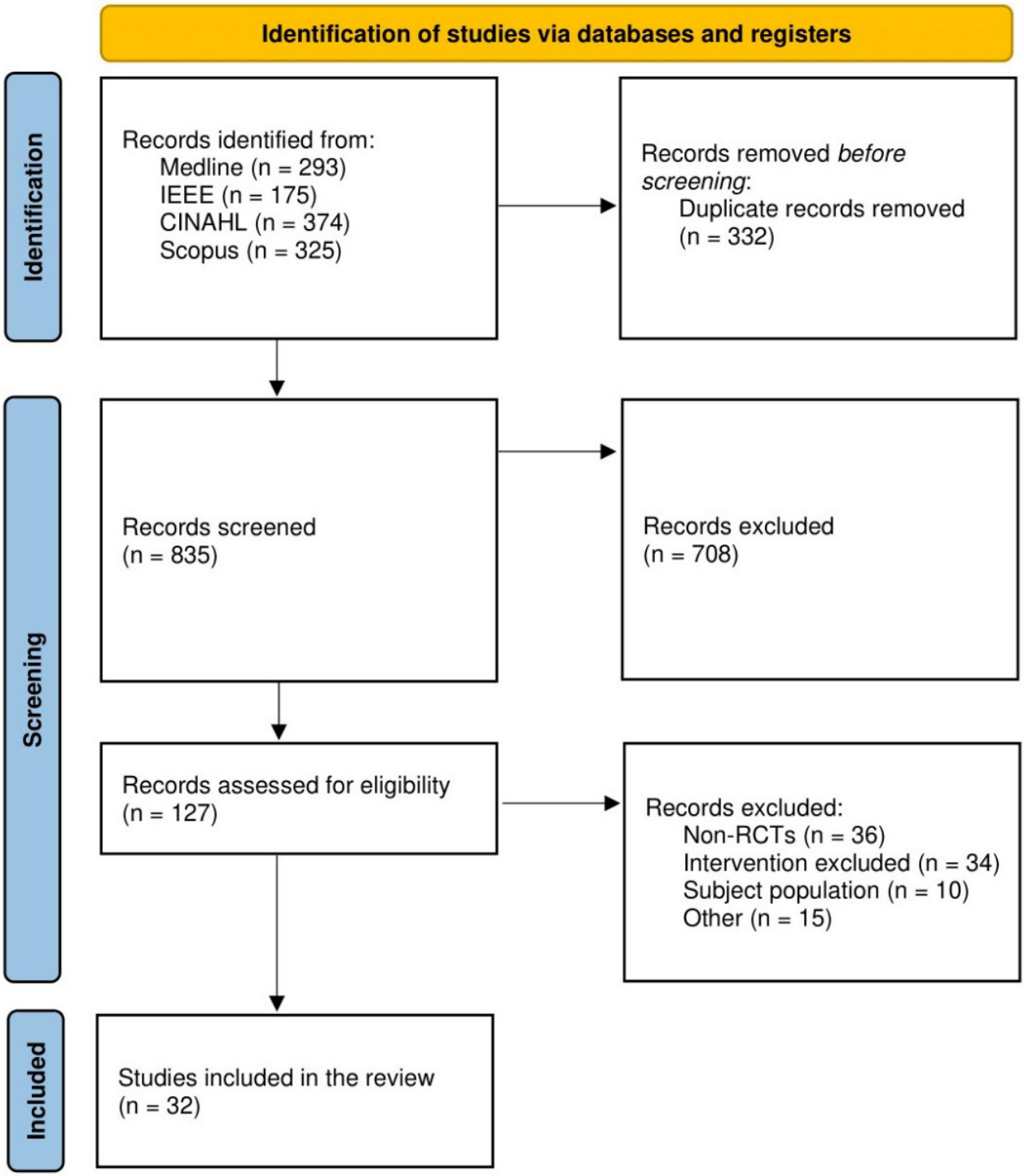
The following study characteristics were extracted: study design, study population, country, intervention type, control and comparison intervention, mean age of participants in both arms, percentage of female participants, years since diagnosis, length of intervention in weeks, number of sessions, length of time to final follow-up assessment in weeks, attrition in the intervention and control arms upon conclusion of the intervention, and primary outcome measures. In studies where there were multiple intervention or control arms, attrition data from intervention arms were combined and attrition data from control arms were combined for meta-analysis purposes [7].

Each study was scored on the basis of its technological elements by WB in accordance with the published manuscript. The scoring system used was adapted from Barakat et al [8], who developed it for use in web-based eTherapy programs targeted to eating disorders. Studies were scored in accordance with four domains:

- multimedia channels (written text="1," graphics/images="1," audio/voiceover="2," video="3," simulation/3D virtual reality="4");
- degree of user interactivity (questionnaires="1," quizzes="1," goal setting/to‐do list="2," homework tasks="3," user dashboard="3," forums="3," self‐monitoring tools="4," interactive exercises="4," virtual games="4," video coaching with professional="4");
- level of automated feedback (motivational pop‐ups="1," reminders="2," nonpersonalized feedback="3"), personalized feedback (telephone or email)= "4");
- technological device through which the program was made accessible (outdated technology; eg, compact disk read‐only memory (CD‐ROM)="1," modern technology; eg, tablet, desktop computer, laptop, and mobile phone="2").

The reporting of technological features of digital health interventions was often vague or absent in publications; therefore, it was necessary to survey each author regarding their examined web-based intervention to clarify scores in accordance with the system adapted from Barakat et al [8]. An email was sent out to authors who had not completed the survey. These authors were sent a summary of the initial scoring by WB and asked to confirm that all the elements listed were present in their study and to identify any that were missed. In total, 19 of the 32 studies responded and the remaining 11 provided additional elements that were not identified in the initial screen by WB. In total, 3 of the 33 studies [9-11] provided access to the program itself and provided confirmation of elements.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram [6]. RCT: randomized controlled trial.



Quality Appraisal

The methodological quality of the final studies included in the review was assessed by author WB using the Cochrane Risk of Bias tool 2 [12]. The tool uses five domains to assess the bias within study design: allocation concealment and random sequence generation, blinding of participants and study personnel, blinding of outcome assessment, incomplete outcome data, and selective reporting of data. Briefly, each domain is composed of questions, to which the response can be "yes," "possibly yes," "possibly no," "no," and "no information." Based on answers, each domain is scored as "low risk," "some concerns," or "high risk," whereby the overall judgement is made by the combination of all domains.

Statistical Analyses

Analyses were conducted using Stata (Stata/SE 16 for Mac, StataCorp).

Meta-analysis

The Stata command *meta* [13] was used for meta-analysis. The input variables required by *meta* were contained in a 2×2 table; that is, number of individuals who did and did not experience the "outcome event" in either the treatment or the control group.

To compare study attrition rates between the intervention groups and the control group, risk ratios were meta-analyzed using the default restricted maximum likelihood method. Log-transformed relative risk of attrition was computed separately for each study, weighted by the inverse of study variance, and pooled to create a summary effect. Studies with no loss to follow-up in one of the intervention or control arms were adjusted by adding 0.5 to each cell within a study's 2×2 table, enabled by meta-analysis settings within Stata. Studies with no attrition in both intervention and control arms were excluded from the meta-analysis as is best practice [14]. Heterogeneity was assessed using the I^2 statistic. Publication bias was assessed by funnel plots. Data were exponentiated and also presented as a Forest plot using the *meta forestplot* command in Stata.

Meta-regression

Meta-regression was used to examine the relationship between the log-transformed relative attrition rates, and various study characteristics, participant demographics, and intervention technological elements. Study characteristics included the type of control, type of intervention, duration of the study, and duration to the final follow-up. Type of control was dummy-coded as follows: wait-list, active control, usual care or active control computerized (1, 2, 3, 4). The type of intervention was dichotomized owing to small numbers of observations dummy-coded as follows: containing an exercise or physiotherapy condition or other (1, 2). The length of the intervention period was converted to years and evaluated as a continuous variable.

Demographic elements included the type of MS, mean age of the participants, female-to-male ratio, and years since disease onset. The type of MS was dummy-coded as either all MS types or relapsing-remitting MS (1, 2).

The intervention technological elements that could support engagement were as follows: multimedia channels, interactivity, and feedback; overall score was the sum of these 3 elements. This analysis omitted the element that assessed how the intervention was made accessible since all interventions used desktop computers. An attempt was made to further clarify whether interventions were designed to accommodate multiple platforms, but this remained unclear for many studies.

Variables were first examined individually and then jointly as a single meta-regression model informed by a directed acyclic graph.

Results

Results Overview

The database search retrieved 1167 articles, of which 187 were duplicates. After duplicates were removed, titles and abstracts of the remaining 835 were screened. Of these, 127 articles were obtained for full-text review. In total, 95 papers did not meet the inclusion criteria and were excluded after full-text review. Of the excluded papers, 34 were not applicable interventions, 36 were not RCTs, 10 had the wrong subject population, and 15 others were excluded for other reasons (Multimedia Appendix 1). In total, 32 papers that described 32 studies were eligible for meta-analysis.

Study Characteristics

Details of the study characteristics for the 32 included papers are reported in Table 1. Of the 32 studies, 10 were published before 2015, and the earliest paper was published in 2004. A total of 23 of 32 studies included all types of MS, 8 included only people with relapsing-remitting MS, and 1 included primary-progressive MS (PPMS). Two included any type of MS if fatigue was a symptom, 2 others included any type of MS if depression was a symptom, 1 assessed whether participants experienced migraines, and 1 only assessed whether participants experienced severe disability. The mean length of interventions was 14.0 (SD 10.2) weeks, and the mean length to last follow-up was 20.9 (SD 12.1) weeks. In total, 20 of 32 studies reported a high risk of bias upon quality appraisal with none reporting low risk (Figure 2).

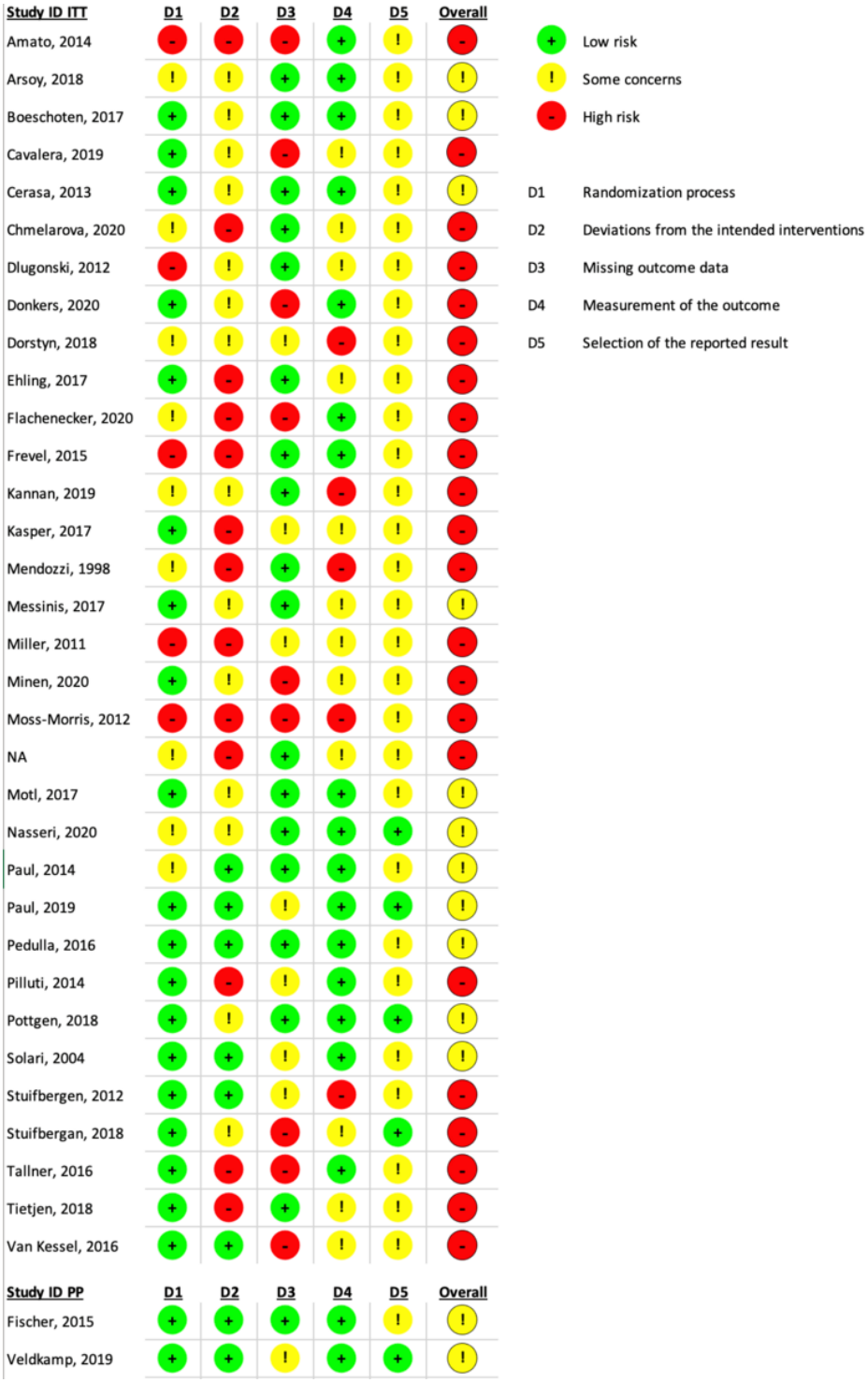
Table 1. Characteristics of the included studies.

Study	Population	Country	Participant age (years), mean	Females, %	Intervention type	Control type	Length of the intervention (weeks)	Maximum follow-up (weeks)
Dlugonski, 2012 [15]	Relapsing-remitting multiple sclerosis	United States	46.65	84	Exercise	Wait-list	12	13
Donkers, 2020 [16]	All types of multiple sclerosis	Canada	41.30	85	Exercise	Wait-list	4	4
Ehling, 2017 [17]	All types of multiple sclerosis	Austria	48.60	45	Exercise	Active	24	26
Flachenecker, 2020 [18]	All types of multiple sclerosis	Germany	47.00	61	Exercise	Usual care	13	39
Frevel, 2015 [19]	Relapsing-remitting multiple sclerosis	Germany	45.60	84	Exercise and falls prevention training	Active	12	12
Kannan, 2019 [20]	All types of multiple sclerosis	United States	55.75	70	Exercise	Wait-list	8	22
Motl, 2011 [21]	Relapsing-remitting multiple sclerosis	United States	45.85	91	Exercise	Wait-list	13	13
Motl, 2017 [22]	All types of multiple sclerosis	United States	51.90	85	Exercise	Wait-list	24	24
Nasseri, 2020 [23]	Primary-progressive multiple sclerosis	Germany	51.10	51	Exercise	Usual care	12	12
Pilluti, 2014 [24]	All types of multiple sclerosis	United States	49.00	76	Exercise	Active (computer)	8	26
Tallner, 2016 [25]	All types of multiple sclerosis	Germany	40.80	75	Exercise	Wait-list	13	26
Paul, 2014 [10]	All types of multiple sclerosis	United Kingdom	51.25	75	Physiotherapy	Wait-list	12	13
Paul, 2019 [11]	All types of multiple sclerosis	United Kingdom	56.05	77	Physiotherapy	Active	26	39
Amato, 2014 [26]	Relapsing-remitting multiple sclerosis	Italy	40.9	78	Cognitive rehabilitation	Active (computer)	36	49
Boeschoten, 2017 [27]	All types of multiple sclerosis with depression	Netherlands	48.90	82	Problem-solving treatment for depression	Wait-list	10	17.4
Chmelařová, 2020 [28]	All types of multiple sclerosis	Czech Republic	41.90	78	Cognitive rehabilitation	Active	8	8
Fischer, 2015 [29]	All types of multiple sclerosis with depression	Germany	45.28	78	Cognitive behavioural therapy	Wait-list	9	26
Messinis, 2017 [30]	Relapsing-remitting multiple sclerosis	Greece	45.60	69	Cognitive rehabilitation	Usual care	10	26
Minen, 2020 [31]	All types of multiple sclerosis with migraine	United States	39.70	89	Cognitive behavioural therapy	Usual care	13	26
Moss-Morris, 2012 [32]	All types of multiple sclerosis	United Kingdom	40.95	0.82	Cognitive behavioural therapy	Wait-list	8	10
Pedulla, 2016 [33]	All types of multiple sclerosis	Italy	47.60	68	Cognitive training	Active (computer)	8	26

Study	Population	Country	Participant age (years), mean	Females, %	Intervention type	Control type	Length of the intervention (weeks)	Maximum follow-up (weeks)
Pottgen, 2018 [34]	All types of multiple sclerosis with fatigue	Germany	41.35	81	Cognitive behavioural therapy	Wait-list	12	24
Solari, 2004 [35]	All types of multiple sclerosis	Italy	43.70	64	Cognitive behavioural therapy	Active (computer)	8	16
Stuifbergen 2012 [36]	All types of multiple sclerosis	United States	N/A[a]	89	Group cognitive rehabilitation	Wait-list	8	13
Stuifbergen 2018 [37]	All types of multiple sclerosis	United States	N/A	88	Group cognitive rehabilitation	Usual care	8	26
van Kessel, 2016 [38]	All types of multiple sclerosis with fatigue	New Zealand	43.00	74	Cognitive behavioural therapy	Active	10	10
Veldkamp, 2019 [39]	All types of multiple sclerosis	Belgium	52.4	58	Cognitive training	Active	8	12
Cavalera, 2019 [40]	Relapsing-remitting multiple sclerosis	Italy	42.73	65	Mindfulness training	Active	8	26
Kasper, 2017 [41]	All types of multiple sclerosis	Germany	40.1	71	Health literacy	Active (computer)	0.14	0.14
Miller, 2011 [42]	All types of multiple sclerosis	United States	48.10	79	Falls prevention training	Active	52	52
Dorstyn, 2018 [9]	Relapsing-remitting multiple sclerosis	Australia	41.30	85	PowerPoint to build work skills	Wait-list	4	4
Cerasa, 2013 [43]	Relapsing-remitting multiple sclerosis	Italy	32.7	75	Cognitive rehabilitation	Active (computer)	6	6

[a]N/A: not applicable.

Figure 2. Risk of bias assessment summary using the Cochrane Risk of Bias tool. ITT: intention-to-treat; PP: per-protocol.



Study Populations

The mean sample size across all 32 studies was 97.7 (SD 121.8). The smallest study contained 18 participants [19] and the largest 682 participants [41]. The median numbers of participants in the intervention and control arms at the final follow-up were 33 and 27, respectively.

Types of Digital Health Intervention and Control

All studies included a program aimed at people with MS delivered via web-based technology, with 2 studies [36,37] that included in-person components. The program ranged from 0.14 week to 52 weeks intended duration. In total, 15 of the 32 studies included a cognitive behavioral training element, 11 included an exercise training component, 2 included web-based physiotherapy, 1 included a fall prevention intervention, 1 included a program to build "returning to work" skills, 1 included a program to build health literacy, and 1 included a mindfulness intervention.

In total, 12 of the 32 studies used a wait-list control, while 7 studies used a usual care control. Eight studies used a noncomputerized active control, which included a physical activity regime [17-19,42], physiotherapy [11], modified version

of the intervention [38,39], or non–MS-specific psychoeducation [40]. Five studies used a computerized active control, which were all broadly modified versions of the intervention [26,33,35,41,43].

Digital Health Intervention Characteristics App

Complete scoring information for each individual study is available in Multimedia Appendix 2.

Multimedia

The mean multimedia score across all 32 studies was 3.5 (SD 2.2). In total, 29 of the 32 studies contained written text. Two of the remaining 3 studies that did not include written text were a gamified web-based cognitive training program [36,37], and 1 study described a web-based, asynchronous messaging system that uses electronic personal health records [42]. In total, 15 of 32 studies contained video or animation, half of all studies including images or graphics, and only one study used audio voiceover to augment written material [27].

Interactivity

The mean interactivity score across all 32 studies was 5.2 (SD 3.3). Eight of 32 studies contained homework tasks [9,27,32,34,36-38,40], which varied from set questions or self-directed learning to diary entries between course sessions. In total, 15 of 33 studies contained a self-monitoring component [10,11,15,16,18,21-25,31,32,38,42], which varied from a validated tool integrated within the program to generic mood-related questions and diary entries. Nine studies included interactive exercises or games [26,28,30,33,35-37,39,43], 5 studies contained goal-setting elements [15,21,22,24,38], 3 contained forums for interaction with other participants [20-22], 3 contained video coaching or interactive meditation by professionals [15,24,40], and 1 contained a questionnaire or quiz [32].

Feedback

The mean feedback score across all 32 studies was 2.8 (SD 2.4). In total, 14 of 32 studies included feedback [9-11,16,18,19,21,22,27,28,36,37,39,40], which ranged from personalized feedback from a trained staff member or professional to general feedback delivered automatically via the web-based platform. Reminders were included in 6 of 17 studies [15,17,20,25,32,42], which prompted users to complete the sessions, while 2 studies also included reminders on when new content was added or existing content was updated [21,22].

Meta-analysis

The final meta-analysis included 29 studies after 3 exclusions due to no attrition in both the intervention and control arms. Meta-analysis described an overall effect size of =–0.993 (95% CI 0.95-1.04) and a *P* value of .75, which indicates no statistical difference in the degree of attrition between the intervention and control arms (Figure 3). The test of homogeneity of study-specific effect sizes (*Q*) was rejected (*P*<.01). The data also suggested the presence of a medium level of heterogeneity between studies (I^2=50.50). There is evidence of publication bias as described by the funnel plot showing some asymmetry (Figure 4).

Figure 3. Individual effect sizes and forest plot of the difference in the degree of attrition between the intervention and control arms of the included studies within our analysis (negative values indicate greater attrition in the intervention arm than in the control arm and vice versa).

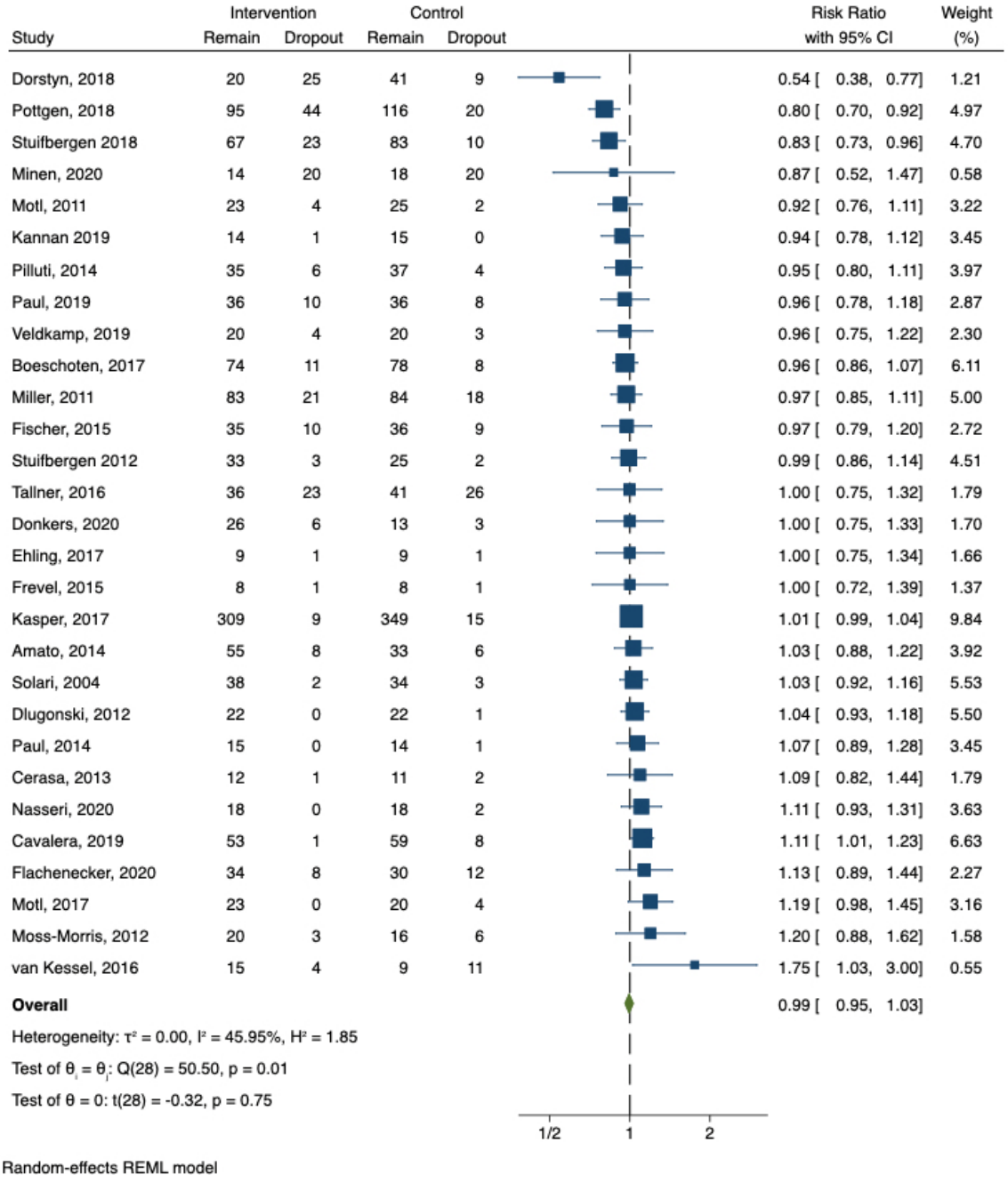
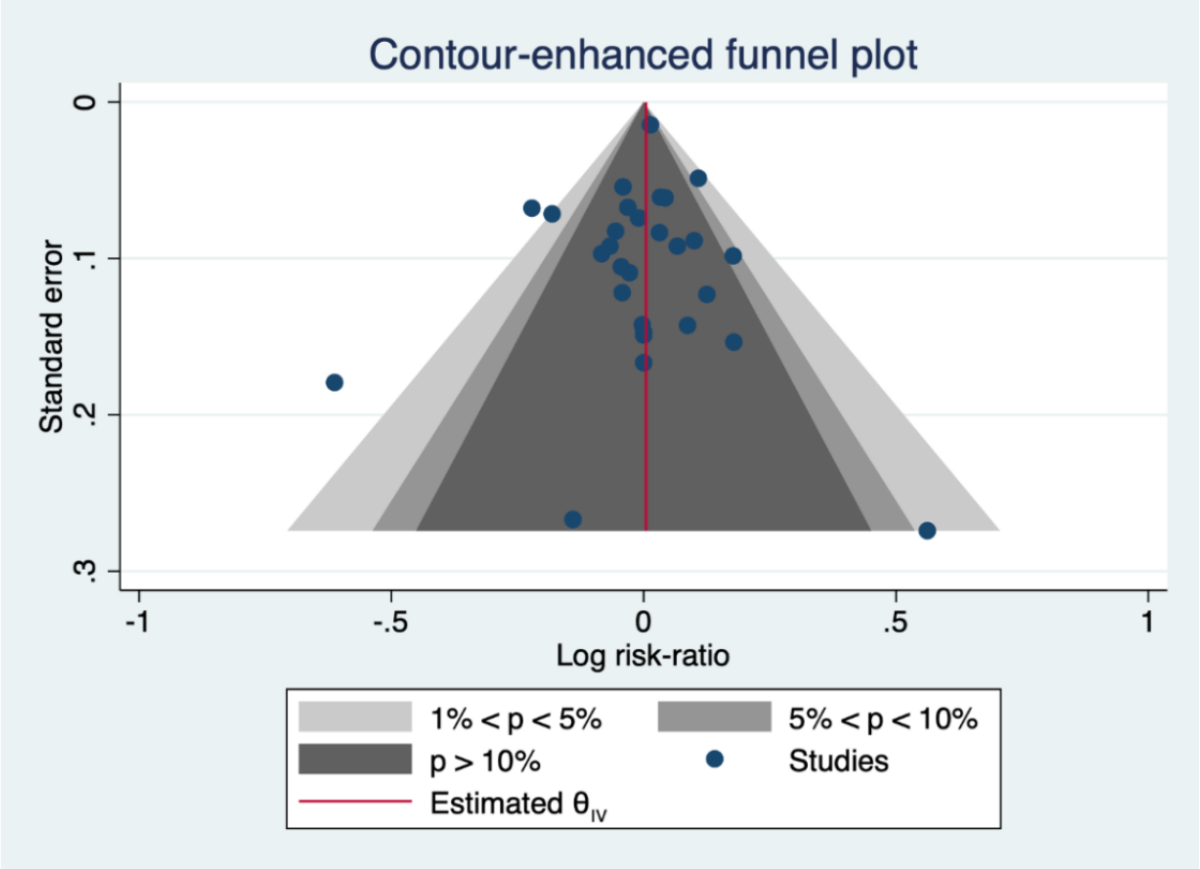


Figure 4. Contour-enhanced funnel plot of the relative attrition rates (>0 indicates a higher attrition rate in the intervention condition).



Meta-regression

Univariable Association

Univariable assessment of study characteristics and eHealth elements are described in Table 2. No significant association between length-of-treatment, weeks to last follow-up, type of control, MS type, feedback or interactivity, or attrition was observed. Lower attrition was observed when eHealth elements overall score and multimedia score were higher.

Table 2. Univariable meta-regression model describing the association between differential attrition and study characteristics.

Variables	Coefficient (95% CI)	*P* value
Length of treatment (years)	1.00 (1.00-1.00)	.88
Active control	1.03 (0.92-1.15)	.64
Usual care	1.07 (0.92-1.23)	.38
Active control computerized	1.06 (0.94-1.20)	.29
Relapsing-remitting multiple sclerosis	1.03 (0.93-1.14)	.60
Primary-progressive multiple sclerosis	1.12 (0.89-1.42)	.31
Non–cognitive behavioral therapy intervention	1.01 (0.92-1.10)	.88
Mean age (years)	1.00 (0.99-1.01)	.60
Female-to-male ratio	0.99 (0.99-1.00)	.01
Years since onset	1.01 (0.99-1.04)	.28
Overall score	1.00 (1.00-1.01)	.20
Multimedia subscore	*1.02* (*1.00-1.04*)[a]	*.03*
Interactivity subscore	1.00 (0.99-1.02)	.47
Feedback subscore	1.00 (0.98-1.02)	.79

[a]Italicized values are significant at *P*<.05.

Multivariable Models

Adjustment for other covariates did materially impact the association of overall score with attrition, which was significant (*P*<.01; Table 3). After adjustment, the multimedia subscore remained a significant association (*P*=.03; Table 4).

Table 3. Multivariable meta-regression model describing the association between overall score and attrition.

Variables	Risk ratio (95% CI)	*P* value
Overall score	*1.01 (1.00-1.03)*[a]	*.01*
Length of treatment (years)	1.08 (0.85-1.39)	.50
Active control	1.04 (0.88-1.22)	.46
Usual care	0.96 (0.79-1.16)	.64
Active control computerized	1.15 (0.97-1.37)	.09
Nonexercise intervention	0.97 (0.88-1.08)	.59
Relapsing-remitting multiple sclerosis	0.99 (0.88-1.12)	.84
Primary-progressive multiple sclerosis	1.11 (0.82-1.49)	.48
Female-to-male ratio	0.99 (0.99-1.00)	.13
Mean age (years)	1.00 (0.99-1.01)	.91

[a]Italicized values are significant at *P*<.05.

Table 4. Multivariable meta-regression model describing the association between multimedia subscore and attrition.

Variables	Risk ratio (95% CI)	*P* value
Multimedia subscore	*1.04 (1.00-1.08)* [a]	*.03*
Length of treatment (years)	1.14 (0.87-1.48)	.32
Active control	1.05 (0.89-1.24)	.53
Usual care	0.99 (0.82-1.20)	.94
Active control computerized	1.09 (0.93-1.28)	.28
Nonexercise intervention	1.01 (0.92-1.11)	.84
Relapsing-remitting multiple sclerosis	1.01 (0.90-1.14)	.80
Primary-progressive multiple sclerosis	1.06 (0.79-1.42)	.68
Female-to-male ratio	1.00 (0.99-1.00)	.34
Mean age (years)	1.00 (0.98-1.01)	.55

[a]Italicized values are significant at *P*<.05.

Discussion

Principal Findings

The aim of this systematic review and meta-analysis was to describe the study characteristics and technological components of digital health interventions that supported people with MS through a program of learning and practicing self-management, and then investigate any associations these factors may have with attrition. Our meta-analysis found no difference in attrition between participants allocated to intervention and control arms of the included RCTs. A multivariable association was found between degree of attrition and overall score, which was not observed in the univariable model, and between the degree of attrition and the multimedia subscore in both univariable and multivariable models. This suggests that among included studies, the interactivity subscore was the main driver in the association observed between overall score and degree of attrition; that is, the likelihood of attrition within the intervention arm compared with the control arm was associated with the inclusion of interactive elements such as self-directed tasks or self-monitoring.

Only 8 of the 32 studies described a design process or framework by which the intervention was developed. Development methods included engagement with the MS community [9,10,31,32,38], design with practitioners and allied health professionals [34], both [31], or followed a preset development plan [23]. There was no association observed between the number or type of technological elements used within the intervention and whether the study used or described a development framework. It is possible that a greater degree of consultation with the MS community during design phases does not necessarily lead to a great quantity of or more complex technological features. This highlights that while engagement with communities during the design phase is important, it is crucial to ensure this engagement is constructive and not perfunctory. Population preferences for number and type of

features require further exploration; however, our findings suggest that elaborate features are not necessarily what people with MS want in a web-based intervention of this kind.

Interestingly, length of intervention was not associated with attrition in either control or intervention arms. Intuitively, the longer the intervention itself, the greater the potential attrition from these studies, in which a negative relationship between attrition and length of treatment for self-management interventions for people with MS has been reported [45]. It is possible that the role of treatment length in attrition may only be true at longer time frames than in the included studies of both meta-analyses. Another likely variable is the number of sessions or time in contact with participants. It is possible that rather than the length of the intervention being a mediator, it is the time participants are exposed to the intervention ("dose"). Studies infrequently report these data; therefore, this needs to be addressed in future studies.

Another interesting finding is that the mean average attrition rates in the intervention and control arms were 14.7% and 15.6%, respectively, which is a useful figure for calculating sample sizes for future studies. A previous meta-analysis investigating attrition rates based on published studies of existing self-management interventions for people with MS reported pooled attrition rates of 16.8% and 14.4% for the intervention and control groups, respectively [45]. Interestingly, studies that used face-to-face delivery of interventions were described as having higher attrition rates. Our analysis supports their pooled data attrition rates, which indicates that there may not actually be a difference between face-to-face and non–face-to-face digital health interventions of this kind. Further, their meta-analysis described associations between attrition and sex, age, and length of intervention, which was not observed in our data set. As their data was pooled between face-to-face and other modes of delivery, it is possible that the web-based mode of delivery may be mediating that relationship in our analysis.

Importantly, 9 of 32 studies did not report any outcomes related to the use of the intervention or technological elements within their intervention; they only reported primary outcome measures related to health outcomes. In total, 15 of 32 studies reported logins, app use, or sessions completed, while only 4 studies reported any interactions between participants and technological elements that were not directly related to the primary outcome measure. This has implications of measuring the "dose" for digital health interventions, whereby factors not reported may have an impact on the outcome measures reported as well as on attrition. Studies describing digital health interventions should address this in the future.

Limitations

Technological elements were underreported in published papers as evidenced by the authors of 10 of 32 studies needing to be followed up to provide information not accessible within the papers. It is possible that this analysis failed to capture the entirety of the included technological elements present within these studies; however, this is unlikely with the follow-up procedure. The need for better reporting of digital health interventions is crucial in ensuring reproducibility and assessment of web-based interventions, especially as several accessible frameworks exist directed at digital tools [44,46] or more generally [47]. Providing access to explore the intervention itself—as 3 authors did for this study—may be another solution, although it is not feasible for some studies.

Digital health interventions are not new; however, there have been fewer published studies on their use in MS compared with other chronic conditions; a greater number of studies would make the meta-regression more reliable. Further stratification of participants may provide greater insight into the attrition in these interventions by accounting for digital literacy, level of access to technology infrastructure, and better general reporting of the intervention design overall.

While the scoring system used for technological component ratings in this review has previously been used and was created in accordance with guidelines for digital intervention development, this system is not yet validated. Several additions were made to this scoring system to accommodate features that were not included, and it is possible that these additions have compromised internal validity. Analysis was undertaken using the previous, unmodified scoring system and revealed no changes to the results using our modified scoring system; however, caution must be applied to interpreting these results before validation has occurred. Additionally, while the scoring was derived from both publication and authors themselves, it is possible there were inaccuracies due to misinterpretation or inability to recall.

It is possible that using qualitative methods and direct observations to understand the ways in which people with MS engage with technological components of these digital health interventions is a better predictor of attrition than merely quantifying and scoring the components themselves. Unfortunately, no studies provided the data required to carry out these analyses. It also possible that analyzing individual components of the subscores for multimedia, interactivity and feedback, rather than overall subscores themselves, may elucidate further associations; however, the weighting within these subscores of the different components aims to address this.

The inclusion criteria were also strict on the definition of the digital health interventions of interest to us. This excluded studies that were exclusively virtual reality interventions and those that were exclusively telemonitoring or conferencing. These exclusions may have biased our data. It is noteworthy that the category of digital health interventions of interest to us is not well-defined in the major taxonomy of digital health interventions published by the World Health Organization; the category *Interventions for Clients*, this describes "communication" but overlooks the structuring and scaffolding of interventions that would be better described as "education" [48].

Conclusions

In conclusion, this study describes no difference in the rates of attrition between participants allocated to the intervention and control arms in digital health interventions for people with MS. An association between overall technological elements score

and attrition was observed; however, the underlying mechanism is unclear. Future studies should further investigate the roles these elements play in retention in web-based interventions.

Acknowledgments

We would like to thank Cecily Gilbert for their assistance in refining our search strategy.

Authors' Contributions

WB, TW, GJ, KG, and SSY contributed to study design. WB contributed to literature search and in data collection and synthesis. WB and SSY were responsible for data analysis. All authors contributed to manuscript development and approval.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Numbers of inclusions/exclusions at each stage of study selection, including reasons for exclusion.
[DOCX File , 18 KB - jmir_v24i2e27735_app1.docx]

Multimedia Appendix 2
Rubric for scoring of individual studies and their technological elements.
[DOCX File , 23 KB - jmir_v24i2e27735_app2.docx]

References

1. GBD 2016 Multiple Sclerosis Collaborators. Global, regional, and national burden of multiple sclerosis 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 2019 Mar;18(3):269-285 [FREE Full text] [doi: 10.1016/S1474-4422(18)30443-5] [Medline: 30679040]
2. Brigo F, Lattanzi S, Bragazzi N, Nardone R, Moccia M, Lavorgna L. Why do people search Wikipedia for information on multiple sclerosis? Mult Scler Relat Disord 2018 Feb;20:210-214. [doi: 10.1016/j.msard.2018.02.001] [Medline: 29428464]
3. Lejbkowicz I, Paperna T, Stein N, Dishon S, Miller A. Internet usage by patients with multiple sclerosis: implications to participatory medicine and personalized healthcare. Mult Scler Int 2010;2010:640749 [FREE Full text] [doi: 10.1155/2010/640749] [Medline: 22096625]
4. Marrie RA, Leung S, Tyry T, Cutter GR, Fox R, Salter A. Use of eHealth and mHealth technology by persons with multiple sclerosis. Mult Scler Relat Disord 2019 Jan;27:13-19. [doi: 10.1016/j.msard.2018.09.036] [Medline: 30296732]
5. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. NPJ Digit Med 2020;3:21 [FREE Full text] [doi: 10.1038/s41746-020-0224-8] [Medline: 32128451]
6. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
7. McKenzie JE, Brennan SE, Ryan RE, Thomson HJ, Johnston RV, Thomas J. Defining the criteria for including studies and how they will be grouped for the synthesis. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. Cochrane Handbook for Systematic Reviews of Interventions (2nd edition). Hoboken, NJ: John Wiley & Sons; 2019.
8. Barakat S, Maguire S, Smith KE, Mason TB, Crosby RD, Touyz S. Evaluating the role of digital intervention design in treatment outcomes and adherence to eTherapy programs for eating disorders: A systematic review and meta-analysis. Int J Eat Disord 2019 Oct;52(10):1077-1094. [doi: 10.1002/eat.23131] [Medline: 31328815]
9. Dorstyn D, Roberts R, Murphy G, Kneebone I, Craig A, Migliorini C. Online resource to promote vocational interests among job seekers with multiple sclerosis: a randomized controlled trial in Australia. Arch Phys Med Rehabil 2018 Feb;99(2):272-280. [doi: 10.1016/j.apmr.2017.08.475] [Medline: 28928026]
10. Paul L, Coulter EH, Miller L, McFadyen A, Dorfman J, Mattison PGG. Web-based physiotherapy for people moderately affected with Multiple Sclerosis; quantitative and qualitative data from a randomized, controlled pilot study. Clin Rehabil 2014 Sep;28(9):924-935. [doi: 10.1177/0269215514527995] [Medline: 24691218]
11. Paul L, Renfrew L, Freeman J, Murray H, Weller B, Mattison P, et al. Web-based physiotherapy for people affected by multiple sclerosis: a single blind, randomized controlled feasibility study. Clin Rehabil 2019 Mar;33(3):473-484. [doi: 10.1177/0269215518817080] [Medline: 30514108]
12. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898. [doi: 10.1136/bmj.l4898] [Medline: 31462531]

13. Bradburn MJ, Deeks JJ, Altman DG. metan-an alternative meta-analysis command. Stata Technical Bulletin 1999;8(44) [FREE Full text]
14. Deeks JJ, Higgins JP, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. Cochrane Handbook for Systematic Reviews of Interventions (2nd edition). Hoboken, NJ: John Wiley & Sons; 2019.
15. Dlugonski D, Motl RW, Mohr DC, Sandroff BM. Internet-delivered behavioral intervention to increase physical activity in persons with multiple sclerosis: sustainability and secondary outcomes. Psychol Health Med 2012;17(6):636-651. [doi: 10.1080/13548506.2011.652640] [Medline: 22313192]
16. Donkers S, Nickel D, Paul L, Wiegers SR, Knox KB. Adherence to physiotherapy-guided web-based exercise for persons with moderate-to-severe multiple sclerosis: a randomized controlled pilot study. Int J MS Care 2020;22(5):208-214 [FREE Full text] [doi: 10.7224/1537-2073.2019-048] [Medline: 33177956]
17. Ehling R, Edlinger M, Hermann K, Dröge K, Seidinger Y, Miller U, et al. Successful long-term management of spasticity in patients with multiple sclerosis using a software application (APP): A pilot study. Mult Scler Relat Disord 2017 Oct;17:15-21. [doi: 10.1016/j.msard.2017.06.013] [Medline: 29055449]
18. Flachenecker P, Bures AK, Gawlik A, Weiland AC, Kuld S, Gusowski K, et al. Efficacy of an internet-based program to promote physical activity and exercise after inpatient rehabilitation in persons with multiple sclerosis: a randomized, single-blind, controlled study. Int J Environ Res Public Health 2020 Jun 24;17(12):4544 [FREE Full text] [doi: 10.3390/ijerph17124544] [Medline: 32599767]
19. Frevel D, Mäurer M. Internet-based home training is capable to improve balance in multiple sclerosis: a randomized controlled trial. Eur J Phys Rehabil Med 2015 Feb;51(1):23-30 [FREE Full text] [Medline: 24755773]
20. Kannan M, Hildebrand A, Hugos CL, Chahine R, Cutter G, Cameron MH. Evaluation of a web-based fall prevention program among people with multiple sclerosis. Mult Scler Relat Disord 2019 Jun;31:151-156 [FREE Full text] [doi: 10.1016/j.msard.2019.04.015] [Medline: 31004969]
21. Motl RW, Dlugonski D, Wójcicki TR, McAuley E, Mohr DC. Internet intervention for increasing physical activity in persons with multiple sclerosis. Mult Scler 2011 Jan;17(1):116-128. [doi: 10.1177/1352458510383148] [Medline: 20921239]
22. Motl RW, Hubbard EA, Bollaert RE, Adamson BC, Kinnett-Hopkins D, Balto JM, et al. Randomized controlled trial of an e-learning designed behavioral intervention for increasing physical activity behavior in multiple sclerosis. Mult Scler J Exp Transl Clin 2017;3(4):2055217317734886 [FREE Full text] [doi: 10.1177/2055217317734886] [Medline: 29051831]
23. Nasseri N, Ghezelbash E, Zhai Y, Patra S, Riemann-Lorenz K, Heesen C, et al. Feasibility of a smartphone app to enhance physical activity in progressive MS: a pilot randomized controlled pilot trial over three months. PeerJ 2020;8:e9303 [FREE Full text] [doi: 10.7717/peerj.9303] [Medline: 32612882]
24. Pilutti LA, Dlugonski D, Sandroff BM, Klaren RE, Motl RW. Internet-delivered lifestyle physical activity intervention improves body composition in multiple sclerosis: preliminary evidence from a randomized controlled trial. Arch Phys Med Rehabil 2014 Jul;95(7):1283-1288. [doi: 10.1016/j.apmr.2014.03.015] [Medline: 24699237]
25. Tallner A, Streber R, Hentschke C, Morgott M, Geidl W, Mäurer M, et al. Internet-supported physical exercise training for persons with multiple sclerosis-a randomised, controlled study. Int J Mol Sci 2016 Sep 30;17(10):1667 [FREE Full text] [doi: 10.3390/ijms17101667] [Medline: 27706046]
26. Amato M, Goretti B, Viterbo R, Portaccio E, Niccolai C, Hakiki B, et al. Computer-assisted rehabilitation of attention in patients with multiple sclerosis: results of a randomized, double-blind trial. Mult Scler 2014 Jan;20(1):91-98. [doi: 10.1177/1352458513501571] [Medline: 23959713]
27. Boeschoten RE, Dekker J, Uitdehaag BM, Beekman AT, Hoogendoorn AW, Collette EH, et al. Internet-based treatment for depression in multiple sclerosis: A randomized controlled trial. Mult Scler 2017 Jul;23(8):1112-1122. [doi: 10.1177/1352458516671820] [Medline: 27682229]
28. Chmelařová D, Fiala L, Dostál M, Lenz J. Intensive computer-assisted cognitive rehabilitation in persons with multiple sclerosis – results of a 12-week randomized study. Cesk Slov Neurol N 2020 Aug 15;83/116(4):408-415. [doi: 10.14735/amcsnn2020408]
29. Fischer A, Schröder J, Vettorazzi E, Wolf OT, Pöttgen J, Lau S, et al. An online programme to reduce depression in patients with multiple sclerosis: a randomised controlled trial. Lancet Psychiatry 2015 Mar;2(3):217-223. [doi: 10.1016/S2215-0366(14)00049-2] [Medline: 26359900]
30. Messinis L, Nasios G, Kosmidis MH, Zampakis P, Malefaki S, Ntoskou K, et al. Efficacy of a computer-assisted cognitive rehabilitation intervention in relapsing-remitting multiple sclerosis patients: a multicenter randomized controlled trial. Behav Neurol 2017;2017:5919841 [FREE Full text] [doi: 10.1155/2017/5919841] [Medline: 29463950]
31. Minen MT, Schaubhut KB, Morio K. Smartphone based behavioral therapy for pain in multiple sclerosis (MS) patients: A feasibility acceptability randomized controlled study for the treatment of comorbid migraine and ms pain. Mult Scler Relat Disord 2020 Nov;46:102489 [FREE Full text] [doi: 10.1016/j.msard.2020.102489] [Medline: 32950893]
32. Moss-Morris R, McCrone P, Yardley L, van Kessel K, Wills G, Dennison L. A pilot randomised controlled trial of an Internet-based cognitive behavioural therapy self-management programme (MS Invigor8) for multiple sclerosis fatigue. Behav Res Ther 2012 Jun;50(6):415-421. [doi: 10.1016/j.brat.2012.03.001] [Medline: 22516321]

33. Pedullà L, Brichetto G, Tacchino A, Vassallo C, Zaratin P, Battaglia MA, et al. Adaptive vs. non-adaptive cognitive training by means of a personalized App: a randomized trial in people with multiple sclerosis. J Neuroeng Rehabil 2016 Oct 04;13(1):88 [FREE Full text] [doi: 10.1186/s12984-016-0193-y] [Medline: 27716336]
34. Pöttgen J, Moss-Morris R, Wendebourg J, Feddersen L, Lau S, Köpke S, et al. Randomised controlled trial of a self-guided online fatigue intervention in multiple sclerosis. J Neurol Neurosurg Psychiatry 2018 Sep;89(9):970-976. [doi: 10.1136/jnnp-2017-317463] [Medline: 29549193]
35. Solari A, Motta A, Mendozzi L, Pucci E, Forni M, Mancardi G, CRIMS Trial. Computer-aided retraining of memory and attention in people with multiple sclerosis: a randomized, double-blind controlled trial. J Neurol Sci 2004 Jul 15;222(1-2):99-104. [doi: 10.1016/j.jns.2004.04.027] [Medline: 15240203]
36. Stuifbergen AK, Becker H, Perez F, Morison J, Kullberg V, Todd A. A randomized controlled trial of a cognitive rehabilitation intervention for persons with multiple sclerosis. Clin Rehabil 2012 Oct;26(10):882-893 [FREE Full text] [doi: 10.1177/0269215511434997] [Medline: 22301679]
37. Stuifbergen AK, Becker H, Perez F, Morrison J, Brown A, Kullberg V, et al. Computer-assisted cognitive rehabilitation in persons with multiple sclerosis: Results of a multi-site randomized controlled trial with six month follow-up. Disabil Health J 2018 Jul;11(3):427-434 [FREE Full text] [doi: 10.1016/j.dhjo.2018.02.001] [Medline: 29477372]
38. van Kessel K, Wouldes T, Moss-Morris R. A New Zealand pilot randomized controlled trial of a web-based interactive self-management programme (MSInvigor8) with and without email support for the treatment of multiple sclerosis fatigue. Clin Rehabil 2016 May;30(5):454-462. [doi: 10.1177/0269215515584800] [Medline: 25952587]
39. Veldkamp R, Baert I, Kalron A, Tacchino A, D'hooge M, Vanzeir E, et al. Structured cognitive-motor dual task training compared to single mobility training in persons with multiple sclerosis, a multicenter RCT. J Clin Med 2019 Dec 10;8(12):2177 [FREE Full text] [doi: 10.3390/jcm8122177] [Medline: 31835502]
40. Cavalera C, Rovaris M, Mendozzi L, Pugnetti L, Garegnani M, Castelnuovo G, et al. Online meditation training for people with multiple sclerosis: A randomized controlled trial. Mult Scler 2019 Apr;25(4):610-617. [doi: 10.1177/1352458518761187] [Medline: 29485319]
41. Kasper J, van de Roemer A, Pöttgen J, Rahn A, Backhus I, Bay Y, et al. A new graphical format to communicate treatment effects to patients-A web-based randomized controlled trial. Health Expect 2017 Aug;20(4):797-804 [FREE Full text] [doi: 10.1111/hex.12522] [Medline: 27981688]
42. Miller DM, Moore SM, Fox RJ, Atreja A, Fu AZ, Lee J, et al. Web-based self-management for patients with multiple sclerosis: a practical, randomized trial. Telemed J E Health 2011;17(1):5-13 [FREE Full text] [doi: 10.1089/tmj.2010.0133] [Medline: 21214498]
43. Cerasa A, Gioia MC, Valentino P, Nisticò R, Chiriaco C, Pirritano D, et al. Computer-assisted cognitive rehabilitation of attention deficits for multiple sclerosis: a randomized trial with fMRI correlates. Neurorehabil Neural Repair 2013 May;27(4):284-295. [doi: 10.1177/1545968312465194] [Medline: 23192417]
44. Agarwal S, LeFevre AE, Lee J, L'Engle K, Mehl G, Sinha C, WHO mHealth Technical Evidence Review Group. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. BMJ 2016 Mar 17;352:i1174. [doi: 10.1136/bmj.i1174] [Medline: 26988021]
45. Arafah AM, Bouchard V, Mayo NE. Enrolling and keeping participants in multiple sclerosis self-management interventions: a systematic review and meta-analysis. Clin Rehabil 2017 Jun;31(6):809-823. [doi: 10.1177/0269215516658338] [Medline: 27401492]
46. Baker TB, Gustafson DH, Shaw B, Hawkins R, Pingree S, Roberts L, et al. Relevance of CONSORT reporting criteria for research on eHealth interventions. Patient Educ Couns 2010 Dec;81 Suppl:S77-S86 [FREE Full text] [doi: 10.1016/j.pec.2010.07.040] [Medline: 20843621]
47. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014 Mar 07;348:g1687. [doi: 10.1136/bmj.g1687] [Medline: 24609605]
48. Classification of digital health interventions v1. World Health Organization. URL: https://www.who.int/reproductivehealth/publications/mhealth/classification-digital-health-interventions/en/ [accessed 2022-01-12]

Abbreviations

CD-ROM: compact disk read - only memory
MeSH: Medical Subject Headings
MS: multiple sclerosis
PPMS: primary-progressive multiple sclerosis
RCT: randomized controlled trial

Edited by R Kukafka, G Eysenbach; submitted 04.02.21; peer-reviewed by S Kitsiou, Y Han; comments to author 07.03.21; revised version received 06.06.21; accepted 14.10.21; published 09.02.22.

Please cite as:
Bevens W, Weiland T, Gray K, Jelinek G, Neate S, Simpson-Yap S
Attrition Within Digital Health Interventions for People With Multiple Sclerosis: Systematic Review and Meta-analysis
J Med Internet Res 2022;24(2):e27735
URL: https://www.jmir.org/2022/2/e27735
doi:10.2196/27735
PMID:35138262

Review

Digital Health Policy and Programs for Hospital Care in Vietnam: Scoping Review

Duc Minh Tran[1], PMD; C Louise Thwaites[1,2], BSc, MBBS, MD; Jennifer Ilo Van Nuil[1,2], BA, MA, PhD; Jacob McKnight[3], BEng, MSc, DPhil; An Phuoc Luu[1], MPH; Chris Paton[3,4], BMBS, BMedSci, MBA; Vietnam ICU Translational Applications Laboratory (VITAL)[5]

[1]Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam

[2]Centre for Tropical Medicine and Global Health, University of Oxford, Oxford, United Kingdom

[3]Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom

[4]Department of Information Science, University of Otago, Otago, New Zealand

[5]See Acknowledgments

Corresponding Author:
Duc Minh Tran, PMD
Oxford University Clinical Research Unit
764 Vo Van Kiet, Ward 1, District 5
Ho Chi Minh City, 700000
Vietnam
Phone: 84 356574593
Email: ductm@oucru.org

Abstract

Background: There are a host of emergent technologies with the potential to improve hospital care in low- and middle-income countries such as Vietnam. Wearable monitors and artificial intelligence–based decision support systems could be integrated with hospital-based digital health systems such as electronic health records (EHRs) to provide higher level care at a relatively low cost. However, the appropriate and sustainable application of these innovations in low- and middle-income countries requires an understanding of the local government's requirements and regulations such as technology specifications, cybersecurity, data-sharing protocols, and interoperability.

Objective: This scoping review aims to explore the current state of digital health research and the policies that govern the adoption of digital health systems in Vietnamese hospitals.

Methods: We conducted a scoping review using a modification of the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines. PubMed and Web of Science were searched for academic publications, and *Th Vi n Pháp Lu t*, a proprietary database of Vietnamese government documents, and the Vietnam Electronic Health Administration website were searched for government documents. Google Scholar and Google Search were used for snowballing searches. The sources were assessed against predefined eligibility criteria through title, abstract, and full-text screening. Relevant information from the included sources was charted and summarized. The review process was primarily undertaken by one researcher and reviewed by another researcher during each step.

Results: In total, 11 academic publications and 20 government documents were included in this review. Among the academic studies, 5 reported engineering solutions for information systems in hospitals, 2 assessed readiness for EHR implementation, 1 tested physicians' performance before and after using clinical decision support software, 1 reported a national laboratory information management system, and 2 reviewed the health system's capability to implement eHealth and artificial intelligence. Of the 20 government documents, 19 were promulgated from 2013 to 2020. These regulations and guidance cover a wide range of digital health domains, including hospital information management systems, general and interoperability standards, cybersecurity in health organizations, conditions for the provision of health information technology (HIT), electronic health insurance claims, laboratory information systems, HIT maturity, digital health strategies, electronic medical records, EHRs, and eHealth architectural frameworks.

Conclusions: Research about hospital-based digital health systems in Vietnam is very limited, particularly implementation studies. Government regulations and guidance for HIT in health care organizations have been released with increasing frequency since 2013, targeting a variety of information systems such as electronic medical records, EHRs, and laboratory information

systems. In general, these policies were focused on the basic specifications and standards that digital health systems need to meet. More research is needed in the future to guide the implementation of digital health care systems in the Vietnam hospital setting.

(J Med Internet Res 2022;24(2):e32392) doi:10.2196/32392

KEYWORDS

digital health; eHealth; policy; Vietnam; hospital care; data; health; electronic medical records; standards; compulsory; patient ID; administrative information; health insurance ID; mobile phone

Introduction

Digital health systems such as electronic health records (EHRs) and patient administration systems used in hospitals in high-income countries (HICs) have been adopted with the dual aim of increasing the quality of patient care and improving hospital finances through cost reductions and new revenue streams. These systems are commonly introduced in response to major government initiatives, often with significant public funding [1]. Despite major challenges and high-profile failures [2], HICs have now reached the point where secondary use of data from digital health systems can, in some cases, enable hospitals and health care systems to become *learning health systems*, using routinely collected data to facilitate research and quality improvement [3]. In recent years, data from hospital-based digital health systems have been used to develop and implement innovative artificial intelligence (AI) systems for monitoring patients and providing clinical decision support (CDS) to health care providers [4].

Although the adoption of digital health systems in low- and middle-income countries (LMICs) such as Vietnam has largely only taken place within the last decade, these solutions have the potential to support the development of universal health coverage and projects working toward addressing sustainable development goals [5,6]. However, in resource-constrained settings, new health care information technologies are often implemented with insufficient funding, infrastructure, regulations, and computer literacy of the staff who will be using them [7-9]. These challenges may be able to be mitigated through the use of open-source software [8,10], mobile technologies [11], and cloud-based data infrastructure [12]. The adoption of new policies and standards (such as Health Level 7's [HL7] Fast Healthcare Interoperability Resources [13]) may also enable simpler and more effective methods for health information exchange than was available to HICs in previous years [14].

Recent initiatives in Vietnam and other LMICs have sought to exploit the potential of digital technologies such as machine learning and low-cost wearable devices in improving critical care at an affordable cost [15,16]. For these innovations to attain scalability and sustainability, their research and development needs to consider the technical infrastructure and the regulatory frameworks that govern local technology adoption [5,6,15,16]. The Vietnamese government has recognized the role that digital health technologies can play in improving health care and optimizing administration processes [17], and along with building national health databases such as the national EHR system, new digital health policies and guidance have been promulgated by the Vietnam Ministry of Health (MoH) [18,19]. Awareness of the government's current regulations and future directions for digital health and the local research evidence in this area is important for the introduction of these emergent technologies in Vietnamese hospitals. This scoping review aims to map and summarize the academic literature and government policies in the field of hospital-based digital health systems in Vietnam. We have chosen a scoping review methodology as it allowed us to effectively explore and summarize information from a wide range of sources, given the rapidly evolving nature of digital health in Vietnam [20,21].

Methods

We conducted this review using a modification of the *PRISMA-ScR* (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews): Checklist and Explanation* guidelines [22].

Eligibility Criteria

The eligibility criteria are listed in Textbox 1.

Textbox 1. Eligibility criteria.

Inclusion criteria

- Policies and guidance documents from the government that regulate and guide the adoption of digital health systems in Vietnam's public hospitals. Academic publications that address digital health systems in Vietnam's hospitals
- Policies and guidance documents that are functioning or to be mandated
- Documents written in Vietnamese or English

Exclusion criteria

- Policies and guidelines that have been replaced by a newer version
- Policies and studies that only examine technical aspects of the digital solution without discussing its adoption and implementation in clinical settings
- Stand-alone apps or digital solutions that are implemented in the hospital settings but not linked with a particular hospital-based information system such as hospital information management systems and electronic medical records

Information Sources

We conducted searches on PubMed, Web of Science, and Google Scholar to identify academic literature. A structured search string (Multimedia Appendix 1) was built for the search on PubMed and Web of Science, and additional relevant publications identified during the full-text screening were retrieved and screened using Google Scholar.

A search of Vietnamese policies was conducted on the Th Vi n Pháp Lu t database [23]. A total of 10 search queries were built and executed separately (Multimedia Appendix 1). All document titles resulting from each of the search queries were extracted and compiled in a Microsoft Excel spreadsheet, from which all duplicates were removed. The MoH Electronic Health Administration's website was used to search for relevant documents that had not been found in the structured database search. Follow-up visits to the website were done throughout the study and at the end of the data-charting phase to identify newly released documents. Relevant documents that were referred to in the reviewed text but not found by the 2 former search strategies were searched for using Google.

No restriction on publication type and publication year was set.

Data-Charting Process

A data-charting form was developed based on the research objectives and guidelines from the PRISMA-ScR guidelines. A draft version of the data-charting form was piloted during the full-text screening, from which discussions were made to modify and finalize the charting form.

All the charting tasks were conducted by one researcher (DMT) using a Microsoft Excel spreadsheet. The charted data were then reviewed by another researcher (CP), and follow-up discussions were arranged to resolve any disagreements between the 2 investigators.

For academic literature, the charting form collected information about the publication year, methods, study population, or source of information reviewed (if the publication is a review), the study scope, the digital health domain being studied, the context that the domain was investigating, and key findings. The charting form of government documents sought to extract information about the domain of intervention, policy enactment time, the intended purpose of the policy, and the ministry that mandated the document.

Selection of Sources of Evidence

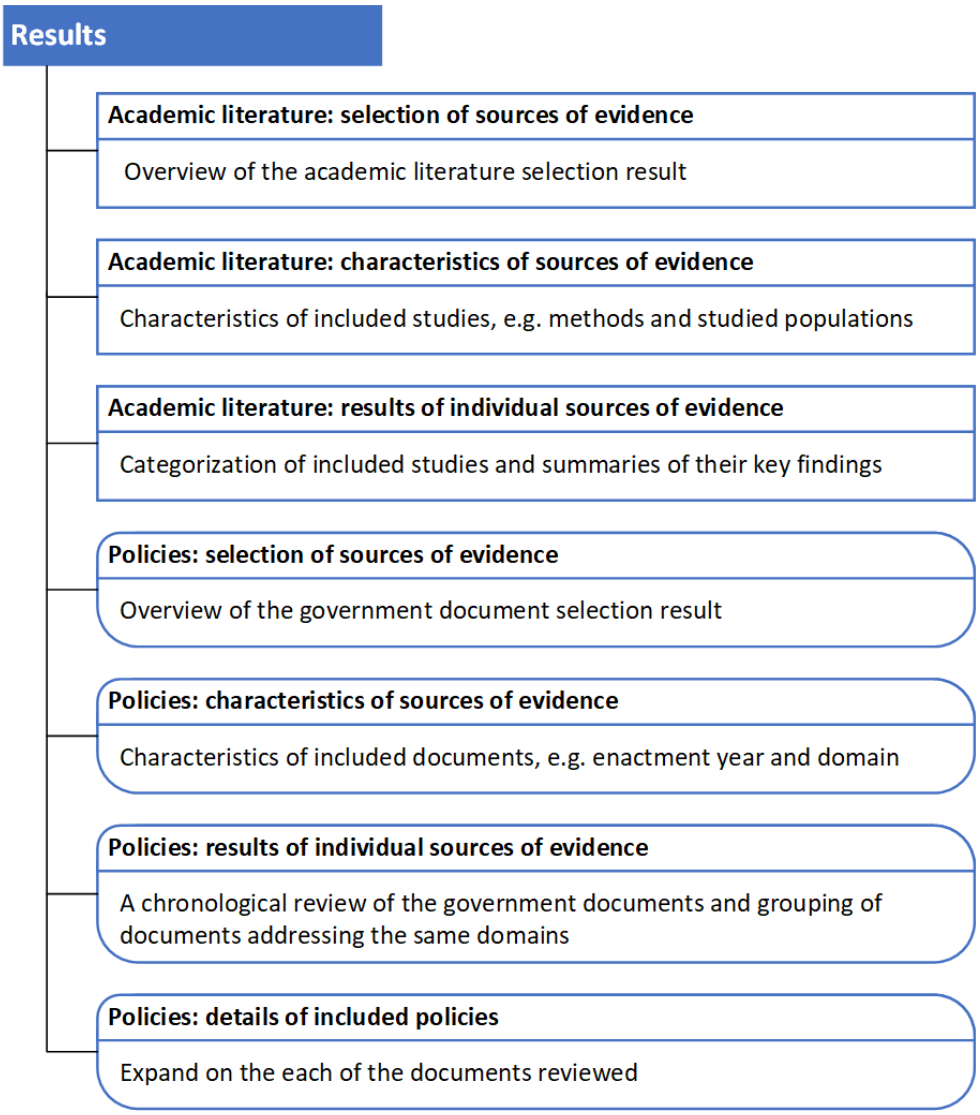
Academic publications were assessed through 3 steps including title, abstract, and full-text screening. Studies that did not meet the eligibility criteria in each step were ruled out from the screening list. Government policies and guidance were selected using a 2-stage process. In the first screening round, the title, introduction, and *scope and regulated entities* section of each document was screened to determine their eligibility. For documents that passed the first screening round, full text was screened in the second round. The selection was based on the eligibility criteria. Any uncertainty during the selection process was discussed between the authors for a final agreement.

Results

Overview

From the initial search results of 2033 academic publications and 266 government documents, 11 academic studies and 20 government documents were included in the review. We have organized the *Results* section into 2 major parts: academic literature and policies. Each part features the source selection process, characteristics of selected sources, and the results of individually selected sources. Figure 1 depicts the layout of the *Results* section.

Figure 1. Layout of the *Results* section. The branches represent the subsections in the *Results* section and the information included therein. Square boxes and round boxes denote the Academic Literature and Policies subsections, respectively.

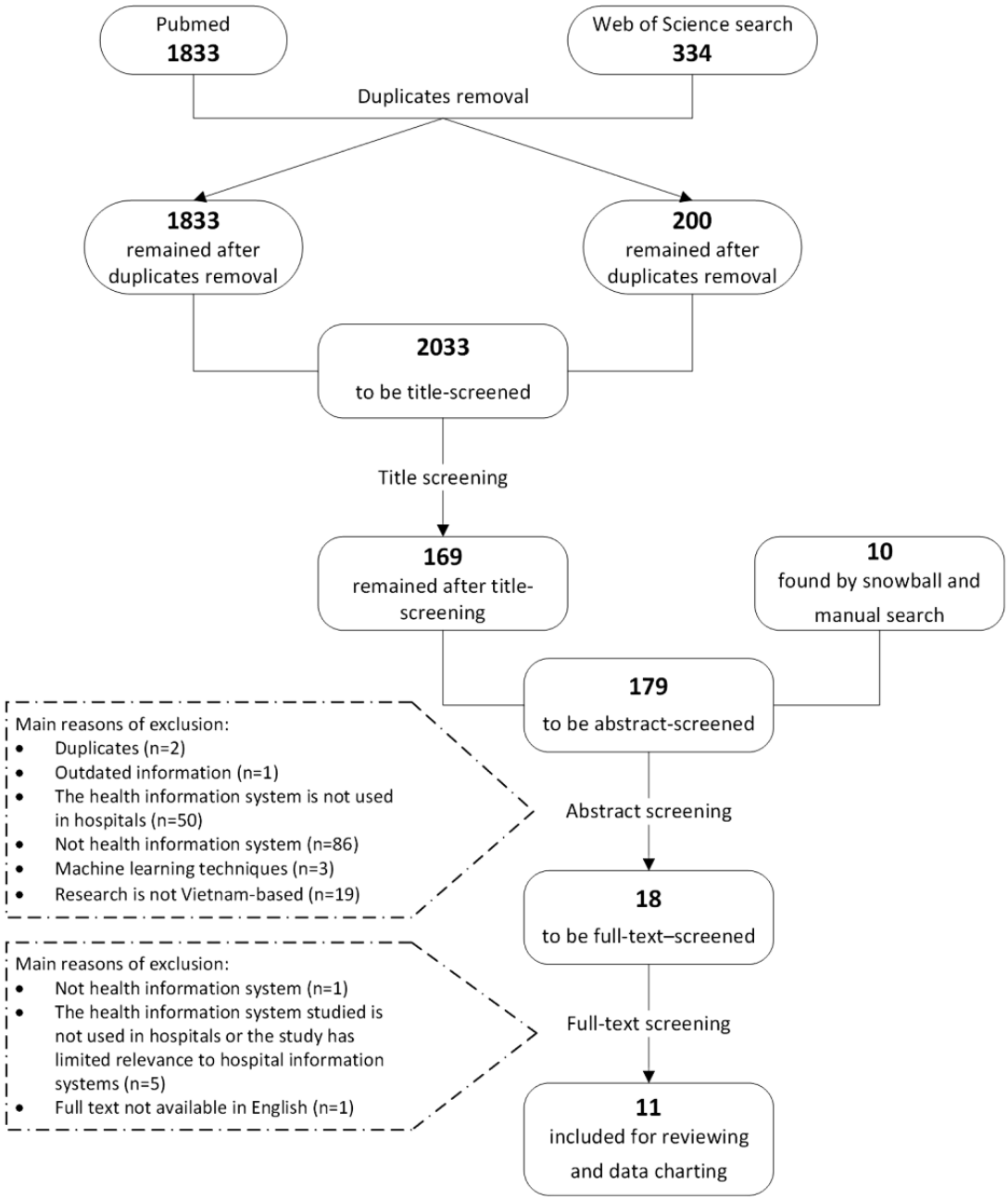


Academic Literature

Selection of Sources of Evidence

The PubMed and Web of Science searches of the academic literature returned 1833 and 334 articles, respectively. After the removal of duplicates, title screening was conducted on 2033 articles (1833 from PubMed and 200 from Web of Science), resulting in 169 marked *considered* for abstract screening. Titles were excluded if they did not imply any relation to the use of information technology (IT) in health care. In addition, the snowballing and manual search found an additional 10 studies, making it a total of 179 abstracts to be screened. The abstract screening ruled out 161 articles that did not meet the eligibility criteria and identified 18 articles eligible for the full-text review. The main reasons for exclusion during abstract screening were as follows: duplicated articles (n=2), the reviewed information was outdated to Vietnam's current digital health situation (n=1), the interventions studied were neither of health information systems (n=86) nor of hospitals' health information systems (n=50), the research was about machine learning techniques (n=3), and the research was not conducted in Vietnam (n=19). Of the full-texts screened, 11 articles were included in the final review. The full-text screening found 7 studies not eligible for the review, of which 1 did not examine health information systems, 5 studies investigated health information systems that were not implemented in hospitals or their findings were of limited relevance to the hospital information systems (HISs), and 1 study's full text was not available in English (Figure 2).

Figure 2. Summary of academic publication selection.



Characteristics of Sources of Evidence

Characteristics of each study including title, authors, publication year, methods, studied population or source of information reviewed, and study scope are summarized in Table 1.

Table 1. Summary of included studies.

Study title	Reference	Methods	Studied population or source of information reviewed	Study scope
Design of Laboratory Information System for Health Care in Vietnam BK-LIS	Vu et al [24]	Case study	Common laboratory test results in Vietnam's hospitals	Described how BK-LIS, a laboratory information system, was designed and developed to support the laboratory activities in Vietnam's hospitals
A Design of Renal Dataflow Control and Patient Record Management System for Renal Department Environment in Vietnam	Vu et al [25]	Case study	Hemodialysis systems in Bach Mai Hospital and E Hospital, 2 central-level hospitals in Vietnam	A case study of the development of BK-HD manager, an IT[a] solution that centrally manages the hemodialysis system in the hospital
Automatic Retrieving Data From Medical Equipment to Create Electronic Medical Records for an e-Hospital Model in Vietnam	Hai et al [26]	Case study	Electronic medical equipment commonly used in Vietnamese hospitals	Demonstrated technical solutions to automatically retrieve data from medical devices. The types of data include images and video data, laboratory test result data, and waveform data
Toward VNUMED for health care research activities in Vietnam	Vo et al [27]	Case study	EMRs[b] from hospitals in Vietnam	Introduced VNUMED, an intermediate database that gathers data from EMRs to support health care research, and related challenges for its development in Vietnam
EMR Visualization for Patient Progress Tracking: Interface Design and System Implementation	VO et al [28]	Case study	Gastroenterologists and EMR data in Thong Nhat central-level hospital	Described the development and testing of EMR visualization, a visualization tool for patient progress tracking, using data from the hospital's EMR system
Strategic Challenges Facing User- and Patient-Centered e-Health in Vietnam	Nguyen et al [29]	Review	IT use in health care in Vietnam before 2012	A review of the IT applications in Vietnam's health sector before 2012. Challenges in implementing patient-centered eHealth in Vietnam were discussed
English-Based Pediatric Emergency Medicine Software Improves Physician Test Performance on Common Pediatric Emergencies: A Multicenter Study in Vietnam	Lin et al [30]	Multicenter, prospective, pretest–posttest study	203 physicians from 11 hospitals across Vietnam	PEMSoft[c], a clinical decision support system, was tested against physicians' performance on a multiple-choice exam
Toward an Electronic Health Record System in Vietnam: A Core Readiness Assessment	Hochwarter et al [31]	Document analysis, participant observation, and in-depth interview	Participant observation and document analysis was conducted in a department of a top-level hospital in Vietnam; in-depth interview with an MoH[d] expert.	Investigation of the Vietnamese health system's core readiness for electronic health record implementation
Open-Source LIMS[e] in Vietnam: The Path Toward Sustainability and Host Country Ownership	Landgraf et al [32]	Reviewing the program reports	Reports from a national LIMS project using an open-source LIMS from 2008 to 2016	Described the building and scale-up of a national LIMS program for clinical and public health laboratories in Vietnam. Outcomes of the program and the lessons learned were discussed. A model for sustainability that could be applied to diverse laboratory programs was proposed
Electronic Health Record Readiness Assessment in Thái Binh Hospital, Vietnam	Nguyen [33]	Cross-sectional study using a scoring tool	Thái Binh provincial hospital	The study assessed the readiness for electronic health record implementation in Thái Binh hospital, Vietnam. The 4 main components of readiness included core readiness, technological readiness, learning readiness, and societal readiness
Artificial Intelligence vs Natural Stupidity: Evaluating AI[f] Readiness for the Vietnamese Medical Information System	Vuong et al [34]	Nonsystematic review	The literature about AI in medicine worldwide, the literature about eHealth in Vietnam, and the Joint Annual Health Review of Vietnam from 2012 to 2016	An overview of AI research and applications in medicine worldwide, proposing a framework to evaluate AI readiness. The assessment of AI readiness in Vietnam's health care sector using the proposed framework

[a]IT: information technology.

[b]EMR: electronic medical record.

[c]PEMSoft: Pediatric Emergency Medicine Software.

[d]MoH: Ministry of Health.

[e]LIMS: laboratory information management system.

[f]AI: artificial intelligence.

Results of Individual Sources of Evidence

The digital health domain that each study addressed, the context in which the domain was considered, year of data collection or reviewed evidence, and key findings of each study are presented in Table 2.

Among the 11 publications reviewed, 5 reported the development and testing of engineering solutions to gather, manage, or visualize data from the medical devices or information systems that were being used in the chosen hospitals. The 3 studies published in 2010 and 2011 described solutions to retrieve and manage data from laboratory devices or hemodialysis systems, whereas the 2 studies published in 2019 involved electronic medical record (EMR)-based solutions.

The other 6 articles addressed aspects of digital health implementation in Vietnam's health care system. Among these, no study examined a specific clinical information system in the hospital setting. EHR implementation was discussed in the 2 readiness assessment studies, in which one interviewed an MoH staff member, whereas the other study surveyed health care workers in a provincial hospital. These stakeholders recognized the benefits offered by EHRs and expressed positive attitudes toward EHR adoption. However, there was a lack of IT infrastructure, basic IT use among the hospital staff such as regular communication via email, and IT training capacity in place for EHR implementation. Lin et al [30] evaluated physicians' performance on a multiple-choice exam with the support of a system (CDS system [CDSS]). The participants improved their exam performance when using the CDS software compared with when not using the CDS software. However, the software was only used for testing purposes and not implemented in the studied hospitals. The national laboratory information management system program published by Landgraf et al [32] was largely used in public health laboratories rather than integrated into a hospital-based system. The 2 reviews by Nguyen et al [29] and Vuong et al [34] provided an overview of how health care facilities in Vietnam had implemented IT applications. Specifically, barriers for the development of patient-centered eHealth and AI in medicine in Vietnam were summarized, including the lack of national health databases, unstandardized data collection, and paper-based workflows. Importantly, most of the data analyzed in these 6 studies were collected in 2016 or earlier. The CDS performance and the EHR readiness assessments were conducted in the period from 2010 to 2014. The paper by Landgraf et al [32] was based on program reports from 2008 to 2016, whereas the review by Nguyen et al [29] provided insights into Vietnam eHealth before 2012. Discussions in the review by Vuong [34] were significantly carried out upon the analysis of the MoH annual reports from 2012 to 2016.

Table 2. Summary of the digital health domains and findings from academic publications.

Study	Digital health domain	Context in which the intervention was considered	Year of data collection or reviewed evidence[a]	Summary of findings relevant to digital health systems in hospitals in Vietnam
Vu et al [24]	Laboratory information system	Common laboratory test results in hospitals in Vietnam	2010 or earlier[b]	• Presented how a laboratory information system was designed to solve the paper-based laboratory result management.
Vu et al [25]	Hemodialysis management system	Hemodialysis systems in Vietnam's hospitals	2010 or earlier[b]	• A central management system for hemodialysis machines in 2 hospitals in Vietnam was designed and tested.
Hai et al [26]	Data acquisition from medical devices	Electronic medical equipment commonly used in Vietnamese hospitals	2011 or earlier[b]	• An engineering solution to automatically retrieve data from medical devices such as ultrasound, ECG[c], and laboratory devices for personal health records.
Vo et al [27]	A database of EMR[d] data	EMRs from hospitals in Vietnam	After 2019	• The lack of standardized EMR use among Vietnamese hospitals posed a challenge for data gathering and research. VNUMED is a database that aims to collect and standardize data from different EMR systems.
VO et al [28]	Data visualization for EMRs	Data in EMR from a Vietnamese hospital	Between 2015 and 2019[b]	• A data visualization app to track patient progress based on data collected from the local EMR system was developed and tested by the gerontologists as the end users. The testing results showed positive feedback from the end users regarding usability. The app was being updated for large-scale testing.
Nguyen et al [29]	Readiness for patient-centered eHealth	Health care in Vietnam before 2012	Before 2012	• Provided an overview of information technology applications in Vietnam's health care system from its beginning until 2011.
Lin et al [30]	CDS[e]	Physicians in hospitals	2010 to 2011	• The study provided evidence that CDS technologies can improve physicians' medical knowledge in the context of Vietnamese hospitals.

Hocwarter et al [31]	EHR readiness	Hospitals in Vietnam	2013	• Provided evidence on the core readiness to a national EHR including the following:
				  • Identification of needs for future changes that will be addressed by the EHR system.
				  • Challenges posed by the status quo that demanded an EHR system: the existing system of medical records; quality of the existing record-keeping practice; the numbering system for medical records; patient identification methods in medical records; use of daily admissions and discharge lists; medical record archiving after patient discharge; medical record preservation when in archive; practice of ICD-10[f] and use of ICD-10 in reality.
				  • Planning for the new EHR project by the Ministry of Health.
				• Integration of technology:
				  • Integration with the current services; plan to integrate the EHR system with the existing hospital information systems; use of health informatics standards; the use of defined interfaces or gateways in data exchange.

Landgraf et al [32]	Laboratory information system	Mainly district health centers and public health laboratories in provinces with a high prevalence of HIV	2008 to 2016	• Described the development, deployment, and operation of a national LIMS[g] project using an open-source LIMS. • Proposed factors for the sustainability of a health information system in Vietnam: (1) selection of appropriate technology, (2) capacity building and knowledge transfer, (3) financial viability, (4) leadership and management, and (5) alignment with national health strategies.

Study	Digital health domain	Context in which the intervention was considered	Year of data collection or reviewed evidence[a]	Summary of findings relevant to digital health systems in hospitals in Vietnam

Nguyen et al [33]	EHR readiness	A provincial hospital in Vietnam	2013 to 2014	• Provided evidence on EHR readiness in a Vietnamese provincial hospital, including the following:
				  • Core readiness: needs for change, planning, suitability of infrastructure, and integration of new technology with the existing services in the hospital.
				  • Technological readiness: need for information technology adoption and information technology infrastructure's capability to implement an EHR.
				  • Learning readiness: information technology training for hospital and implementation staff.
				• Societal readiness:
				  • Electronic communication with other organizations, data exchange between organizations, and sociocultural elements between health workers and patients.

Vuong et al [34]	Readiness for AI[h] in medicine	Health care of Vietnam	2012-2018 (mainly 2012-2016)	• Readiness for AI in medicine in Vietnam was assessed based on 3 factors: financial support, technological, and sociopolitical. In general, although AI in medicine research and political commitment was somewhat promising, the technical factor was seen as weak and inadequate.

[a]The period when primary data are collected or reviewed sources are published.

[b]On the basis of published year and reported grant period.

[c]ECG: electrocardiogram.

[d]EMR: electronic medical record.

[e]CDS: clinical decision support.

[f]ICD-10: International Classification of Diseases, Tenth Revision.

[g]LIMS: laboratory information management system.

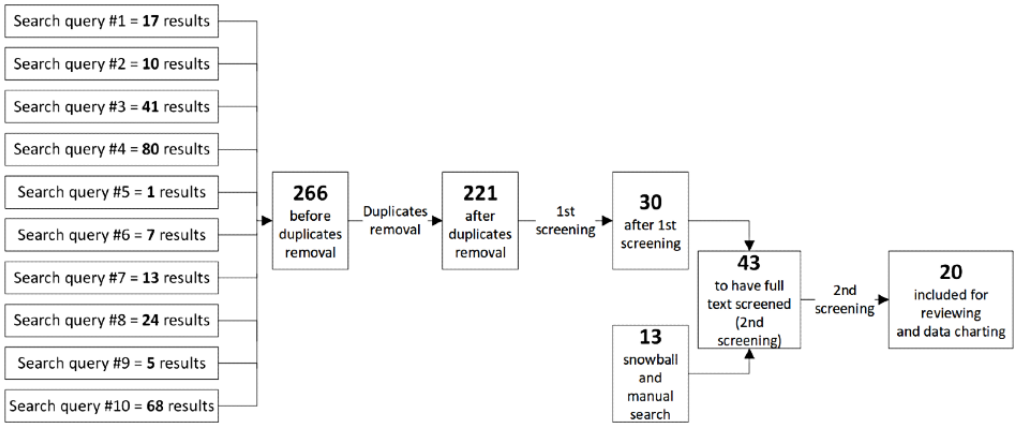
[h]AI: artificial intelligence.

Policies

Selection of Sources of Evidence

A total of 10 search queries on the *Th Vi n Pháp Lu t* database returned 266 results. After removing duplicate titles, there were 221 results remaining in the first screening round. In total, 30 articles were selected from this screening phase with 191 articles excluded for the following reasons: unrelated to HIS adoption (n=157), limited relevance to HIS adoption (n=13), not a policy or guidance document (n=5), expired (n=13), and documents whose contents are covered in an included policy (n=3; Figure 3). Manual scanning on the MoH's Electronic Health Administration website and snowballing from the reviewed documents helped identify 13 potential policies and guidance documents. Full text was screened for a total of 43 documents, which yielded 20 documents including for review and data charting.

Figure 3. Summary of government document selection.



Characteristics of Sources of Evidence

Table 3 presents an overview of the policies and guidance documents included. Characteristics of the documents were charted, including the enactment time points, policy ID numbers and titles, the digital health domains related to the policies, and the ministries that mandated the policies. Owing to the lengthy original titles, shorter denoted titles were created for easier reference of the documents in this paper.

Table 3. Summary of policy documents.

Valid from	Policy ID number and title	Denoted title	Domain	Ministry
December 2006	Decision 5573/QD-BYT year 2006 on Guideline for Hospital Information Management Systems [35]	The HIMS[a] Guidance	Hospital information management system	MoH[b]
June 2013	Decision 2035/QD-BYT year 2013 on Terminology Systems and Data Exchange Standards Recommended for Health IT[c] [36]	The Recommended Standards for HIT[d]	Standards for hospital information systems	MoH
October 2014	Decision 4159/QD-BYT year 2014 on Guidance on Ensuring Security of Electronic Health Data in the Health Sector [37]	The Cybersecurity Guidance	Cybersecurity in health organizations	MoH
March 2015 or January 2017[e]	Circular 53/2014/TT-BYT on Required Conditions for Provision of Health IT Activities [38]	The Required Conditions for HIT	Conditions for provision of health IT	MoH
October 2015	Decision 4495/QD-BYT year 2015 on Guideline for Developing Local Information Safety and Security Rules in Health Facilities [39]	The Guidance for Local Cybersecurity Policy	Cybersecurity in health organizations	MoH
October 2015	Decision 4494/QD-BYT year 2015 on Response Procedures for Information Safety and Security Issues in the Health Sector [40]	The Guidance for Cybersecurity Response	Cybersecurity in health organizations	MoH
September 2016	Decision 5004/QD-BYT year 2016 on The Architectural Framework of the Social Health Insurance Information System [41]	The Social Health Insurance EAF[f]	Electronic health insurance claim	MoH
June 2016	Decision 917/QD-BHXH year 2016 on Announcement of the Health Insurance Portal version 2 [42]	The Insurance Portal V2 Guidance	Electronic health insurance claim	Vietnam Social Security
September 2017	Decision 4210/QD-BYT year 2017 on Requirements for Standard and Format of Output Data Used in Management, Assessment and Reimbursement of Insurance-Paid Health Care Expenses [43]	The Claim Standardization Guidance	Electronic health insurance claim	MoH
August 2017	Decision 3725/QD-BYT year 2017 on Guidelines for Functionalities, Interoperability, Infrastructure and Human Resources for Establishing and Implementing Laboratory Information Systems at Healthcare Facilities [44]	The LIS[g] Guidance	LIS	MoH
February 2018	Circular 54/2017/TT-BYT on Assessment Criteria for Information Technology Implementation in Healthcare Facilities [45]	The HIT Maturity Model	HIT maturity	MoH
July 2018	Circular 39/2017/TT-BTTTT on Technical Standards for IT Implementation in State Organizations [46] (replaced circular 22/2013/TT-BTTTT)	The Recommended Standards for IT in State Organizations	Standards for IT applications in state organizations	MIC[h]
March 2018	Circular 48/2017/TT-BYT on Regulations on Data Exchange in Management and Reimbursement of Health Insurance Claims [47]	The Electronic Claim Regulations	Electronic health insurance claim	MoH
December 2018	Decision 7603/QD-BYT year 2018 on The Service Coding System for Healthcare Management and Health Insurance Reimbursement version 6 [48]	The Terminology and Service Coding System version 6	Electronic health insurance claim	MoH
October 2019	Decision 4888/QD-BYT year 2019 on The Smart Health IT Implementation and Development Scheme from 2019 to 2025 [49]	The Smart HIT Scheme	Digital health strategies	MoH
March 2019	Circular 46/2018/TT-BYT on Regulations for Electronic Medical Records [50]	The Regulations for EMRs[i]	EMRs	MoH
November 2019	Decision 5349/QD-BYT year 2019 on Implementation Plan for Electronic Health Record [51]	The EHR[j] Plan	EHRs	MoH
December 2019	Decision 6085/QD-BYT year 2019 on the eHealth Architectural Framework version 2.0 [52]	The EHAF[k] version 2	Architectural framework	MoH
December 2020	Decision 5316/QD-BYT year 2020 on The Digital Transformation in Health Care Scheme Until 2025 and Navigated Toward 2030 [53]	The Digital Transformation Scheme	Digital health strategies	MoH
May 2020	Decision 2153/QD-BYT year 2020 on Regulations on Creation, Utilization and Management of Health ID [54]	The Health ID Regulation	National health ID	MoH

[a]HIMS: Hospital Information Management System.

[b]MoH: Ministry of Health.

[c]IT: information technology.

[d]HIT: health information technology.

[e]March 2015 or January 2017: applicable since March 2015 for organizations that had not implemented any health information systems and since January 2017 for organizations that had implemented health information systems.

[f]EAF: electronic architecture framework.

[g]LIS: laboratory information system.

[h]MIC: Ministry of Information and Communications.

[i]EMR: electronic medical record.

[j]EHR: electronic health record.

[k]EHAF: eHealth architectural framework.

Results of Individual Sources of Evidence

A total of 20 government documents were combined into 11 groups based on the digital health domains or subjects that they addressed. These groups included hospital information management systems, general and interoperability standards, cybersecurity in health organizations, conditions for the provision of health IT, electronic health insurance claims, laboratory information systems (LISs), health IT (HIT) maturity, digital health strategies, EMRs, EHRs, and eHealth architectural framework (EHAF). Table 4 presents these policy groups and the main purpose of each policy. In the following section, we have summarized the policies and their objectives in chronological order.

Most of the digital health policies reviewed were released and came into effect after 2013, excluding the Hospital Information Management System (HIMS) Guidance, which was enacted in 2006. The HIMS Guidance was an instruction regarding the functionalities and standards for HIMS used in state health facilities [35]. In 2013, the MoH promulgated guidance for interoperability standards and terminology systems applicable for health information systems [36]. Some of these are mandatory such as the HL7 messaging version 2 or 3, Digital Imaging and Communications in Medicine, SDMX-HD (Statistical Data and Metadata Exchange–Health Domain), and International Classification of Diseases, Tenth Revision (ICD-10) Clinical Modification, whereas the others are recommended, including the HL7 Clinical Document Architecture, HL7 Continuity of Care Document, World Health Organization Anatomical Therapeutic Chemical, and Logical Observation Identifiers Names and Codes. Standards for general IT applications are guided by the Ministry of Information and Communications (MIC) [46].

The 3 Cybersecurity Guidance documents for state health organizations were published in 2014 and 2015. The Cybersecurity Guidance covers 16 components of an IT system in a health organization such as network, databases, applications, account management, and data transmission [37]. As organizations using IT systems are required to develop and implement their own cybersecurity policy, the Guidance for Local Cybersecurity Policy provides general instructions for the making of these local policies [39]. Finally, the Guidance for Cybersecurity Response is a standard procedure to identify, classify, report, and handle cybersecurity issues in health facilities [40]. In March 2015, the Required Conditions for HIT announced the essential criteria that organizations must meet when implementing HIT in Vietnam [38]. Accordingly, organizations must satisfy specific requirements for IT infrastructure, information security, IT workforce, and implementation of several technologies.

In 2016 and 2017, the MoH and the Vietnam Social Security (VSS) initiated the electronic social health insurance claim system, which was accompanied by the promulgation of several regulations and guidance. First, the Social Health Insurance Electronic Architectural Framework was published, describing the system's structure and operation [41]. A web-based portal was established, on which health providers can submit claims and check insurance information. To provide instructions for health facilities on using the portal, the VSS published the Insurance Portal V2 Guidance [42]. As claim data must be collected and formatted in a standardized manner, the Claim Standardization Guidance is intended to provide the necessary instructions for health facilities on claim preparation [43]. The Vietnam social health insurance system uses a national terminology and service coding system to form billing codes from medical documentation. In 2018, this coding system was in its sixth iteration [48]. Finally, the Electronic Claim Regulations defines the responsibilities of health providers and insurance agencies surrounding claim sharing and feedback [47].

Table 4. Main purposes of the policies.

Group of domains and denoted title	Purpose of the document	Valid from
Hospital information management system		
The HIMS[a] Guidance	A guidance for functionalities and standards of HIMSs	December 2006
General and interoperability standards		
The Recommended Standards for HIT[b]	A list of nomenclature systems and interoperability standards that the MoH[c] required or recommended for health information systems	June 2013
The Recommended Standards for IT[d] in State Organizations	A list of general IT standards that the MIC[e] required or recommended for health information systems	July 2018
Cybersecurity in health organizations		
The Cybersecurity Guidance	A comprehensive guidance of cybersecurity measures for health facilities	October 2014
The Guidance for Local Cybersecurity Policy	A guideline for developing organization information safety and security policies for health facilities	October 2015
The Guidance for Cybersecurity Response	A guidance for classifying, identifying, reporting, and handling cybersecurity issues in health facilities	October 2015
Conditions for provision of health IT		
The Required Conditions for HIT	Criteria that health facilities must satisfy when implementing digital health systems. Four areas addressed in the circular include IT infrastructure, information security, human resource, and specific criteria for some HIT systems	March 2015 or January 2017
eHealth insurance claim		
The Social Health Insurance EAF[f]	Explaining the architecture model for the social health insurance information system to be built by the MoH	September 2016
The Insurance Portal V2 Guidance	Announcing the launching of the Health Insurance Portal version 2 with an installation manual attached	June 2016
The Claim Standardization Guidance	A guidance for standardization of claims data including variable definition and data standards	September 2017
The Electronic Claim Regulations	Responsibilities of the health care organizations and the insurance agencies in electronic claim exchange and investigation	March 2018
The Terminology and Service Coding System version 6	A common list of health services with the relevant codes used in the social health insurance claim and reimbursement, updated to version 6	December 2018
LIS[g]		
The LIS Guidance	A guidance for functionalities and standards of LISs	August 2017
HIT maturity		
The HIT Maturity Model	Seven levels of HIT application applicable for health care organizations, made of 8 key components and capabilities. Criteria for each HIT level and component were provided herein.	February 2018
Digital health strategies		
The Smart HIT Scheme	Presenting the agenda to develop and implement digital and smart technologies in Vietnam's health care for the period from 2019 to 2025	October 2019
The Digital Transformation Scheme	Presenting the agenda to comprehensively implement IT in Vietnam's health care until 2025 and navigated toward 2030	December 2020
EMRs[h]		
The Regulations for EMRs	Criteria for EMR development and implementation to abide by Health Care Law and replace paper medical records	March 2019
EHR[i]		
The EHR plan	The national plan of the MoH to build and implement the EHR system in Vietnam	November 2019
The Health ID Regulation	The health ID system used for eHealth data of Vietnamese residents	May 2020
Architectural framework		

Group of domains and denoted title	Purpose of the document	Valid from
The EHAF[j] version 2	Explaining the architecture model of key IT systems and databases built by the MoH	December 19

[a]HIMS: health information management system.

[b]HIT: health information technology.

[c]MoH: Ministry of Health.

[d]IT: information technology.

[e]MIC: Ministry of Information and Communications.

[f]EAF: electronic architectural framework.

[g]LIS: laboratory information system.

[h]EMR: electronic medical record.

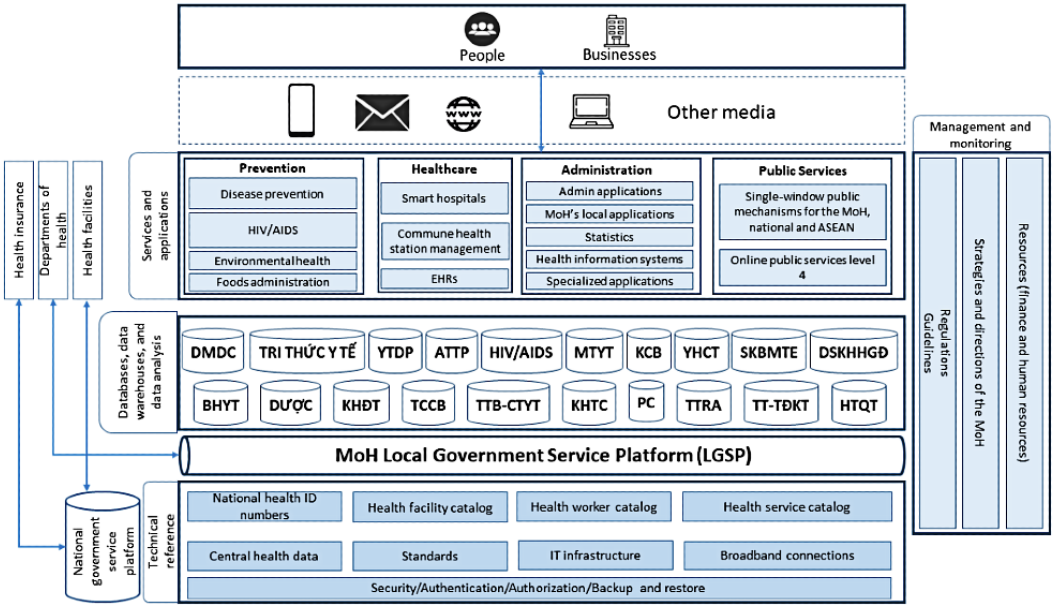
[i]EHR: electronic health record.

[j]EHAF: eHealth architectural framework.

Guidance for LISs was published in August 2017, addressing the functionalities and standards required for such systems in Vietnam [44]. From 2018 to 2020, the MoH issued several policies that govern key digital health systems such as the maturity model for HIT, EMRs, and EHRs. From February 2018, the *HIT Maturity Model* began to be used as a roadmap for health care organizations to build and upgrade their digital health system [45]. Similar to the Healthcare Information and Management Systems Society's Analytics Electronic Medical Record Adoption Model [55], this maturity model features 7 levels of HIT application, each of which is built on a specific set of criteria. The most basic level (level 1) only requires the implementation of HISs to manage outpatient services, whereas the highest level (level 7) defines a paperless environment that deploys a comprehensive range of HIT such as HIS, EMR, LIS, radiology information system and picture archiving and communication systems (RIS-PACS), and CDSS with interoperability between the systems. EMR adoption is governed by the *Regulations for EMRs*, enacted in March 2019 [50]. The circular sets out the requirements for EMRs so that their use abides by Health Care Law and is eligible to completely replace paper medical records. In the same year, the MoH announced a plan to roll out EHRs on a nationwide scale, starting in municipal cities [51]. Although EMRs serve inpatient care and can vary in their specifications depending on the specialty in which they are implemented, the EHR system can be seen as the digital version of the paper-based primary health record [56] and is a lifetime health document of every Vietnamese resident. Information in EHRs is mainly collected and updated by the district health facilities and outpatient clinics. The MoH has also announced the National Health ID system in the Health ID Regulation [54], in which each health ID number is unique and represents an individual for their lifetime.

An overall picture of the IT systems that the MoH has been developing is depicted in their EHAF, presented in Figure 4 (adapted from the Vietnam MoH [52]). The framework was first published in 2015 [57] and subsequently updated in 2018 [58] and 2019 with the EHAF version 2 [52]. According to the EHAF version 2, EMRs and EHRs are the 2 main components of the health care building block. In 2019, the MoH announced the Smart HIT Scheme, presenting an agenda to develop and implement digital and smart technologies in Vietnam's health care for the period from 2019 to 2025 [49]. By the end of 2020, the Digital Transformation Scheme was published, setting the goals to implement IT in multiple areas of the health sector including state administration, cashless payment and telehealth, disease prevention and primary care, and health care [53].

Figure 4. The Ministry of Health's eHealth Architecture Framework version 2.0 (adapted from the Vietnam Ministry of Health [52]). The following are English translations for the abbreviations—ATTP: food safety; BHYT: health insurance; DMDC: reference catalogs; DSKHHGD: population planning; D C: pharmacy; HTQT: international collaboration; KCB: health care; KHĐT: research and training; KHTC: finance; MTYT: environmental health; PC: legislation; SKBMTE: maternal and child health; Tri th c y t : health knowledge; TCCB: human resources; TTB-CTYT: medical devices and infrastructure; TTRA: inspection; TT-TĐKT: reward; YHCT: traditional medicine; YTDP: preventive medicine. EHR: electronic health record; IT: information technology; MoH: Ministry of Health.



Details of Included Policies

In the subsequent sections, we have expanded on the key policies in each digital health domain.

Architectural Framework: The EHAF Version 2

The framework is intended to guide the future developments of eHealth in the public sector, ensuring their compatibility with the current components and directed toward common goals.

The EHAF version 2.0 (decision 6085/QD-BYT year 2019 on the eHealth Architectural Framework Version 2.0 [52]) has 7 layers including the following:

1. End users (people and businesses)
2. Communication channels (means to communicate with applications and services of the MoH such as computers, smartphones, and information portals)
3. Services and applications layer. This layer comprises 4 building blocks including disease prevention, health care, administration, and public services
4. Databases and data analysis tools layer. This includes multiple databases specialized for each health area managed by the MoH, for example, traditional medicine, pharmacy, and health insurance
5. The MoH's Local Government Service Platform. This platform provides shared supporting services for the upper layers. The MoH's Local Government Service Platform is also able to communicate with other ministries and provinces through the National Government Service Platform
6. Technological infrastructure layer
7. Management and monitoring layer

An illustration of the EHAF version 2.0 can be found in Figure 4 (adapted from the Vietnam MoH [52]).

Digital Health Strategies

The Smart HIT Scheme

The Smart HIT Scheme (decision 4888/QD-BYT year 2019 on the Smart HIT Implementation and Development Scheme from 2019 to 2025 [49]) seeks to use digital health, particularly smart technology, in Vietnamese health care. The scheme highlighted the Industrial Revolution 4.0 concept and the technologies characterizing this era, for example, big data, AI, and the internet of things. Following the intention of the Vietnamese government to take advantage of this revolution, the scheme points out the readiness of Vietnam's health system for implementing digital health technologies along with the impact that these technologies can have on the health system. Finally, the scheme presents an agenda to promote and implement digital and smart health IT in Vietnam. The information relevant to the scope of this study is summarized in the subsequent sections.

The scheme aims to *implement and develop digital health and smart health technologies for a modern, high-quality, equitable, efficient, and internationally integrated health system* and *to promote residents' access to health information so that they can use a highly efficient health service and have their health continuously protected, taken care of, and promoted during their lifetime*. The following goals are specified:

1. Developing a smart health care and disease prevention system
2. Promoting IT implementation in health facilities for administrative process improvement and reducing hospital overload: adopting EMRs to replace paper records, using a cashless payment system for hospital billings, and establishing smart hospitals
3. Promoting IT implementation in health administration: installing the electronic office system, public portals, and

single-window information system of administrative procedures, promoting level 3 and level 4 web-based public service, building a smart health administration

To achieve these goals, the scheme proposes 9 areas of action from 2019 to 2025, including the following:

1. Building the regulatory framework, guidance, standards, and economic–technical norms
2. Building the health IT infrastructure
3. Building a smart health care and disease prevention system
4. Building a smart health care system
5. Building a smart health administration system
6. Developing the workforce
7. Promoting smart health IT research, development, and implementation
8. International cooperation
9. Educating the public's awareness of smart health care

In the scope of this study, we presented the detailed agenda of area 1—building the regulatory framework, guidance, standards, and economic–technical norms—and area 4, building a smart health care system (Textboxes 2 and 3).

Textbox 2. Area 1 of the Smart Health Information Technology Scheme: building the regulatory framework, guidance, standards, and economic–technical norms.

Area 1 of the Smart Health Information Technology Scheme

- Building eHealth architecture framework as a prerequisite for information technology implementation in the health system
- Developing regulations for the resident health ID system
- Developing standards for interoperability between health information technology (HIT) systems, that is, health station management software, electronic medical records, and electronic health records (EHRs)
- Developing economic–technical norms for HIT, in which HIT costs are a part of the total health service cost
- Building policy for managing and using electronic nomenclature and coding systems in the health system
- Building policy for using EHRs
- Developing regulations for cybersecurity and privacy protection for health information in the web-based environment
- Building human resource policies for HIT specialists

Textbox 3. Area 4 of the Smart Health Information Technology Scheme: Building a Smart Health Care System.

Area 4 of the Smart Health Information Technology Scheme

- Updating the management software and digital health systems in hospitals:
  - Develop health information systems that adhere to the national and international standards, in which interoperability between the information systems and between the information system and the digital devices (lab devices, imaging diagnosis devices, interactive screens, personal mobile devices, etc) is ensured
  - Standardizing the National Health ID system
  - Building smart hospitals—health care organizations consult the Health Information Technology Maturity Model to develop their smart hospital roadmap
- Electronic medical records (EMRs) are implemented in all health care organizations according to the EMR rollout timeline in circular 46/2018/TT-BYT on the *Regulations for EMRs*, aiming for paperless medical records and cashless payments in hospitals
- Establishing and scaling up information kiosks in hospitals
- Promoting artificial intelligence (AI) application in health care with the following priorities:
  - Building interoperability standards to implement the Internet of Medical Things as an infrastructure to operate clinical decision support systems (CDSSs)
  - Developing real-time CDSSs closely integrated with EMRs
  - Imaging diagnosis assistance
  - Surgery assistance
  - Encouraging businesses and health care organizations to build big data systems embedded with AI algorithm to support clinical decision-making
  - Disease diagnosis, treatment, and prevention with traditional medicines
  - Applying AI in specialties such as imaging diagnosis, cardiovascular diseases, respiratory diseases, orthopedics, cancer, obstetrics, and pediatrics

The Digital Transformation Scheme

The Digital Transformation Scheme (decision 5316/QD-BYT year 2020—The Digital Transformation in Health Care Scheme until 2025 and navigated toward 2030 [53]) seeks to implement IT in all aspects of health care as it defines digital transformation in health care as *the comprehensive application of information technology which prioritizes the cutting-edge digital technology that can make positive changes in all aspects of health care.*

The key 4 areas addressed in this scheme are state administration, cashless payment and telehealth, disease prevention and primary care, and health care. In the health care area, the scheme aims for 15% (210/1400) and 50% (700/1400) of hospitals in the country to successfully adopt paperless EMRs and cashless payment by 2025 and 2030, respectively.

Textbox 4 summarizes the areas and subareas of action set out in the scheme.

Area 5 (digital transformation in primary care and disease prevention) and area 6 (digital transformation in hospitals) are considered priorities over the others. We present details of area 1.2 (Textbox 5) and area 6 in subsequent sections.

Area 6, ie, Digital Transformation in Hospitals is largely similar to area 4 (building a smart health care system) of the Smart HIT Scheme. A comparison between these is shown in Table 5.

Textbox 4. Summary of areas of action in the Digital Transformation Scheme.

Areas of action

- Area 1: infrastructure development
  - Area 1.1: awareness education
  - Area 1.2: building the regulatory framework, guidance, standards, and economic–technical norms
  - Area 1.3: building and upgrading the information technology infrastructure
  - Area 1.4: building health databases
  - Area 1.5: building digital health platforms
  - Area 1.6: ensuring cybersecurity
  - Area 1.7: international cooperation, research, and innovations in digital health
  - Area 1.8: workforce development
- Area 2: implementing information technology in administration and public services
- Area 3: promoting investments in digital health from businesses and hospitals
- Area 4: digitalization in societies
- Area 5: digital transformation in primary care and disease prevention
- Area 6: digital transformation in hospitals

Textbox 5. Area 1.2 of the Digital Transformation Scheme: building the regulatory framework, guidance, standards, and economic–technical norms.

Area 1.2 of the Digital Transformation Scheme

- Publishing regulations and guidance for doing trials of novel digital health products; developing digital health platforms
- Developing regulations and standards for data exchange and interoperability between health information systems based on international standards
- Developing regulations for collecting and managing health data; building a decree for the national health database
- Providing guidance for digital health technologies; updating guidance for building smart hospitals and paperless hospitals
- Developing regulations for protecting security, safety, and confidentiality of health data in the web-based environment
- Developing regulations and guidance for electronic authentication in health care
- Implementing and updating the MoH's e-government structure
- Developing financial mechanisms for health IT services as part of the overall health care service and mechanisms for hiring health IT services
- Developing policies and regulations for telemedicine and e-prescription so that patients can use remote health care services

Table 5. Comparing area 4 (decision 4888/QD-BYT year 2019) and area 6 (decision 5316/QD-BYT year 2020).

Category	Area 4 of the smart HIT[a] scheme	Area 6 of the Digital Transformation Scheme
Updating management software and digital health systems in hospitals	Yes[b]	Yes
Standardizing the National Health ID system	Yes	Yes
Building smart hospitals based on circular 54/2017/TT-BYT	Yes	Yes
EMR[c] implementation based on circular 46/2018/TT-BYT	Yes	Yes
Building information kiosks	Yes	No
Promoting AI[d] application in health care	Yes	Yes
Conducting telehealth and web-based registration based on decision 2628/QD-BYT year 2020	No[e]	Yes
Implementing the national prescription management system in all health care organizations	No	Yes

[a]HIT: health information technology.

[b]The information was addressed in the document.

[c]EMR: electronic medical record.

[d]AI: artificial intelligence.

[e]The information was not addressed in the document.

EMRs: The Regulations for EMRs

EMR refers to the clinical record system used in hospitals' inpatient departments. They are locally managed by health care institutions and different from the EHR system, which is centrally managed on the national database and intended for use by the commune health facilities and outpatient clinics (circular 46/2018/TT-BYT on *Regulations for EMRs* [50]).

This circular is the first official guidance for EMRs in public health care facilities. It ensures that the use of EMRs can uphold the equivalent law and can legitimately replace paper-based medical records. Health care organizations are eligible to discontinue paper-based medical record–keeping if they can satisfy all the criteria in this circular, which concern the following areas:

1. Content specifications
2. EMR creation and update
3. Storage and backup
4. Access right and secondary use
5. Patient identification
6. Digital or electronic signatures
7. Basic functionalities
8. Adherence to interoperability and IT standards
9. Data security and confidentiality
10. Nomenclature system

A total of 4 essential criteria that must be prioritized are as follows:

- Each patient's EMR is provided with an ID number that is unique in the health organization.
- All information that a traditional medical record collects can also be recorded by EMRs.
- Each EMR must have a digital signature of the one who is responsible for the information entered in that record.
- Data stored in an EMR system are protected by Section 2, Article II of Cyber information Security Law [59].

In addition, the circular also provides the criteria for paperless lab test and imaging diagnosis management with LIS and picture archiving and communication system. Full details of this document can be found in Multimedia Appendix 2.

HIT Maturity: The HIT Maturity Model

Overview

The MoH put into effect the *HIT Maturity Model* (circular 54/2017/TT-BYT on Assessment Criteria for IT Implementation in Health Care Facilities [45]) as a roadmap for HIT implementation in health care facilities. Health care organizations are required to evaluate their current HIT maturity using this model and report the evaluation results to the MoH. The baseline results are a requisite for health care organizations to set goals for their next maturity level. An overall HIT maturity level depends on the maturity level of the individual domains, which are as follows:

1. IT infrastructure
2. Administration and operation software
3. HIS
4. RIS-PACS
5. LIS
6. Nonfunctionality standards
7. Security and information safety
8. EMR

In addition, there are specific capabilities that are required for each maturity level. An organization's overall HIT maturity can range from *HIT level 1* to *HIT level7*.

There are 7 levels (1 to 7) applicable for the IT infrastructure domain and the HIS domain, and 2 levels (basic and advanced) applicable for the other domains (administration and operation software, LIS, nonfunctionality standards, security and information safety, RIS-PACS, and EMRs). The 7 levels of HIT maturity are summarized in subsequent sections. An overview of all the 7 levels can be found in Table 6.

Table 6. Levels of health information technology (HIT) maturity.

	Information technology infrastructure	HIS[a]	LIS[b]	RIS-PACS[c]	EMR[d]	Administration and operation software	Security and information safety	Nonfunctionality criteria	Extra capabilities
HIT level 7	Level 7	Level 7	Advanced	Advanced	Advanced	Advanced	Advanced	Advanced	• "Paperless" hospital if all relevant criteria are met • CDSS[e] level 3 supporting doctors' decisions related to treatment protocols and treatment results using suitably customized templates • Data in CDR[f] is analyzed to improve care quality, patient safety, and care efficiency • Clinical data can be readily shared for stakeholders in care coordination based on HL7[g] standards • Continuous reports of hospital services using the data collected
HIT level 6 (smart hospital)	Level 6	Level 6	Advanced	Advanced	Basic	Advanced	Advanced	Advanced	• CDSS level 2 providing: evidence-based warnings for treatment; that is, health and medicine advice, drug information and interaction check, and initial order and prescription violation identification rules • All structured forms; that is, progress notes, consultation notes, problem lists, and discharge summaries are digitalized • Closed-loop management of drugs, using identification technologies to assist drug administration
HIT level 5	Level 5	Level 5	Advanced	Advanced	N/A[h]	Basic	Basic	Basic	• PACS[i] can replace physical films
HIT level 4	Level 4	Level 4	Advanced	Basic	N/A	Basic	Basic	Basic	• PACS allows doctors to access images outside the imaging department • Electronic ordering • Electronic management of inpatient orders
HIT level 3	Level 3	Level 3	Basic	N/A	N/A	Basic	Basic	Basic	• Electronic records having digital vital sign records, nursing notes, medical procedures, and surgical procedures are stored in CDR • CDSS level 1 assisting electronic prescription (new or historic prescription) • Pharmacy information available in the hospital network and supported with CDSS
HIT level 2	Level 2	Level 2	N/A	N/A	N/A	N/A	N/A	N/A	• A CDR consisting of normenclature and coding systems, pharmacy, orders, and test results (if available) • Data in CDR can be shared between stakeholders for care coordination

	Information technology infrastructure	HIS[a]	LIS[b]	RIS-PACS[c]	EMR[d]	Administration and operation software	Security and information safety	Nonfunctionality criteria	Extra capabilities
HIT level 1	Level 1	Level 1	N/A	N/A	N/A	N/A	N/A	N/A	• Patient's information can be accessed electronically

[a]HIS: hospital information system.

[b]LIS: laboratory information system.

[c]RIS-PACS: radiology information system-picture archiving and communication system.

[d]EMR: electronic medical record.

[e]CDSS: clinical decision support system.

[f]CDR: clinical data repository.

[g]HL7: Health Level 7.

[h]N/A: not applicable.

[i]PACS: picture archiving and communication systems.

HIT Level 1

HIT level 1 includes level 1 of IT infrastructure and level 1 of HIS, which aims to provide electronic access to patient data. This IT infrastructure level essentially requires workstation computers, a local area network, and an internet connection while HIS manages outpatient and pharmacy data, and sends claim data to the health insurance claim system. Adoption of MoH's terminology and service coding system is compulsory.

HIT Level 2

HIT level 2 is an upgrade of IT infrastructure and HIS from level 1 to level 2. This will add to the pre-existing system a dedicated server and the laboratory modules that manage laboratory orders and results. HIT level 2 particularly aims at building a clinical data repository (CDR) of pharmacy data, laboratory orders, test results, and the terminology and service coding system. Data in the CDR can be accessed by multiple stakeholders engaging in patient care.

HIT Level 3

HIT level 3 requires an upgrade of IT infrastructure and HIS from level 2 to level 3. Added components include LIS, administrative applications, and cybersecurity solutions at a *basic* level. At this level, more security and storage infrastructures are provided, including firewall devices, security solutions, and specialized storage. CDR capacity is increased to cover vital sign data, nurse notes, and medical procedures. Health care organizations would start implementing CDSS level 1 to assist doctors with drug information via the facility's intranet.

HIT Level 4

HIT level 4 comprises level 4 IT infrastructure, level 4 HIS, advanced LIS, and a basic level of RIS-PACS, administration and operations software, cybersecurity, and nonfunctional standards. Organizations at this HIT level can manage their imaging diagnosis data electronically using RIS-PACS. Although a basic RIS-PACS has yet to discontinue physical film archiving [50], it offers essential functionalities such as integration with imaging diagnosis machines and HISs, digital image viewing, and measurement. An advanced LIS can manage test samples and supplies, integrate with the HIS, and send out alerts for abnormal test values. A storage network (a storage area network or a network-attached storage), a queue management system with display screens, will complement the IT infrastructure. Outpatient orders from doctors can be made electronically, whereas all inpatient orders will be managed in the digital system.

HIT Level 5

HIT level 5 differs from HIT level 4 in level 5 IT infrastructure, level 5 HIS, and *advanced RIS-PACS*. The advanced level of RIS-PACS is able to conduct multi-site consultation, allowing images to be viewed on multiple devices such as laptops and mobile phones. Furthermore, the *Regulations for EMRs* allow paperless image management if health care facilities have *advanced RIS-PACS* with an eligible storage capacity [50]. At level 5, HISs can manage emergency rooms, operating theaters, appointments, and follow-ups and can integrate electronic patient card.

HIT Level 6

Transitioning from HIT level 5 to HIT level 6 requires organizations to implement a *basic* EMR system in addition to upgrading their IT infrastructure, HIS, administration and operation software, security and data safety, CDSS, drug management, and nonfunctional standards to an *advanced* level. Hospitals qualified for HIT level 6 or higher are certified as a *smart hospital*. The basic EMR system provides inpatient medical records and can integrate with other clinical information systems. Mobile devices working on wireless local area network, hospital security cameras, and backup storage systems are needed to meet level 6 IT infrastructure.

Hospitals at HIT level 6 can receive and use data from multiple information systems such as drug information and interactions, treatment protocols, and nutrition. Clinical information can be accessed via mobile devices. In addition, HIT level 6 features several extra capabilities including a level 2 CDSS, a safe medicine management procedure, and increased digitalization of clinical documents. Level 2 CDSSs can warn doctors of drug interactions and potential treatment risks based on the most updated evidence. The medicine management system provides a closed and automatic medicine management procedure using automated dispensing systems and digital identification devices,

for example, barcodes and radio frequency identification. All structured clinical forms such as progress reports, consultation notes, problem lists, and discharge summaries are digitized.

HIT Level 7

At HIT level 7, all 8 domains are at their highest level. Organizations are now equipped with information kiosks and network monitoring systems. A level 7 HIS provides functionalities to manage EMRs, professional procedures, control information kiosks, and support cashless payment. EMR systems at the *advanced* level have dedicated functionalities to manage medical records including increased storage duration, syncing records, and restoring records. Data stored in the EMR system are better protected and can be exchanged under international standards such as HL7. According to the *Regulations for EMRs* [50], an *advanced* EMR system with qualified storage capacity is eligible to replace the paper-based medical record system. Data from CDR are analyzed to improve quality of care, patient safety, and service efficiency. Insights for operational activities can be continuously generated to inform hospital departments such as inpatient, outpatient, and emergency departments. Organizations at this HIT level can exchange data using HL7 standards. Finally, level 3 CDSSs can support a wider range of diagnosis and treatment decisions (Table 6).

Conditions for Provision of Health IT: The Required Conditions for HIT

The Required Conditions for HIT (circular 53/2014/TT-BYT on Required conditions for provision of health IT activities [38]) can be seen as a summary of the criteria that health care facilities need to satisfy when implementing eHealth systems.

This circular announces the conditions and requirements that health care facilities operating in Vietnam must abide by when implementing and operating health IT systems. Health IT activities, as defined in the circular, involve *providing, transferring, collecting, analysing, storing and exchanging health care data through the IT infrastructure*. The conditions specify 4 areas related to health IT implementation: IT infrastructure, information security, human resource, and operational requirements (Multimedia Appendix 3).

General and Interoperability Standards

This group included 2 guidance documents, one from the MIC and one from the MoH. Although the former addressed the general technical standards that IT applications in state organizations need to adopt, the latter revolved around the national and international standards targeted at health information systems.

The Recommended Standards for HIT

In June 2013, the MoH published decision 2035/QD-BYT (year 2013—Terminology systems and data exchange standards recommended for health IT [36]) recognizing the terminology systems and technical standards applicable for health IT systems. Some terminologies and standards are required, whereas others are recommended for adoption in public health care organizations. It is strongly advised in the decision that all public facilities should adopt these systems, and information systems not matching them should plan for relevant transitions. Table 7 presents the standards and nomenclature systems recommended in the decision.

In addition to the standards, the MoH instructs IT systems in the health sector to follow other relevant IT standards from the MIC that were addressed in circular 01/2011/TT-BTTTT, which was then replaced by the new version, circular 22/2013/TT-BTTTT.

Table 7. Standards and nomenclature systems for health information technology systems.

Categories and standard names	Reference	Guidance
Administrative nomenclature		
The list of official administrative units in Vietnam	Decision 124/2004/QD-TTg year 2004 and its amendments	Compulsory
The list of ethnic groups in Vietnam	Decision121-TCTK/PPCD year 2004 and its amendments	Compulsory
The list of occupations in Vietnam	Decision 114/1998/QD-TCTK year 1998 and its amendments	Compulsory
International classification and coding for diseases and medical services		
ICD-10-CM[a]	WHO[b]	Compulsory
ICD-O-3[c]	WHO	Recommended
ICD-10-PCS[d]	WHO	Recommended
ATC[e]	WHO	Recommended
LOINC[f]	Regenstrief Institute	Recommended
International interoperability standards		
Health Level 7 messaging version 2.x or 3.0	Health Level 7	Compulsory
DICOM[g] version 2.0	The National Electrical Manufacturers Association	Compulsory
SDMX-HD[h]	WHO	Compulsory
HL7 CDA[i]	Health Level 7	Recommended
HL7 CCD[j]	Health Level 7	Recommended
ELINCS[k]	The California HealthCare Foundation	Recommended

[a]ICD-10-CM: International Classification of Diseases, Tenth Revision, Clinical Modification.

[b]WHO: World Health Organization.

[c]ICD-O-3: International Classification of Diseases for Oncology, Third edition.

[d]ICD-10-PCS: International Classification of Diseases, Tenth Revision, Procedure Coding System.

[e]ATC: Anatomical Therapeutic Chemical.

[f]LOINC: Logical Observation Identifiers Names and Codes.

[g]DICOM: Digital Imaging and Communications in Medicine.

[h]SDMX-HD: Statistical Data and Metadata eXchange–Health Domain.

[i]HL7 CDA: Health Level 7 Clinical Document Architecture.

[j]HL7 CCD: Health Level 7 Continuity of Care Document.

[k]ELINCS: EHR-Lab Interoperability and Connectivity Specification.

The Recommended Standards for IT in State Organizations

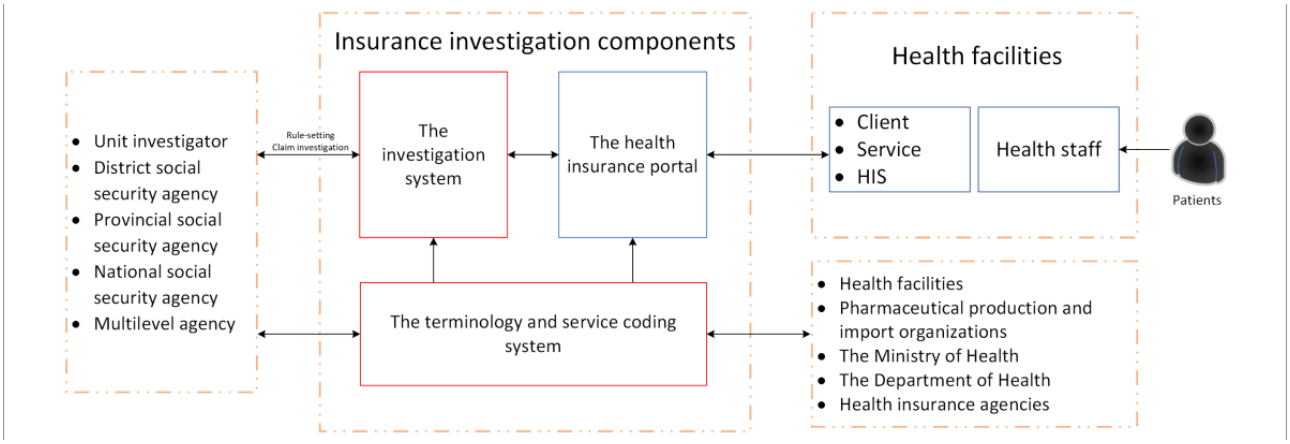
The MIC released circular 39/2017/TT-BTTTT (Technical Standards for IT Implementation in State Organizations [46]) as guidance for the adoption of technical IT standards in state organizations. Its first version is circular 22/2013/TT-BTTTT, released in 2013. The circular advises on 5 categories of standards such as connection, data integration, information access, information security, and public key infrastructure. The MIC encourages systems currently not adopting the standards in this circular to carry out transitions by July 2018. A detailed list of the standards is shown in Multimedia Appendix 4.

Electronic Health Insurance Claim

The Insurance Portal V2 Guidance

Decision 917/QD-BHXH year 2016 (Announcement of the health insurance portal version 2 [42]) from the VSS announces the health insurance portal version 2 and provides detailed guidelines on connection configuration and structuring data sets. The guideline is intended for use by health care facilities and provincial health insurance agencies who are responsible for regular data upload to the portal. The VSS health insurance claim and investigation structure are presented in Figure 5.

Figure 5. The Vietnam Social Security health insurance and investigation structure. HIS: health information system.



The guideline demonstrates 3 protocols that claim data can be submitted, each of which is accompanied by detailed instructions such as setting up connections, properly formatting data, and uploading data to the portal:

- Via web service connection—supports a variety of data submission including clinical data, service validation, monthly reports, health insurance ID checking, examination history, and referral receipt from other facilities.
- Data sync from client software—supports syncing clinical data to the system.
- Direct data entry on the portal—supports clinical data entry to the system.

The portal also offers functionalities for checking health insurance account information, viewing information of previous hospital visits, and checking referral documents between health care facilities.

The Electronic Claim Regulations

The MoH-mandated circular 48/2017/TT-BYT (Regulations on data exchange in management and reimbursement of health insurance claims [47]) as an official regulatory document for web-based health insurance claims and investigation for health care organizations. The circular essentially articulates the responsibilities of health care facilities and the health insurance agency as well as relevant rules in sending data and follow-ups. The fundamental technical requirements previously described in the VSS's guidelines are also outlined in the document (Multimedia Appendix 5).

The Claim Standardization Guidance

The health insurance portal requires claim data to be formatted as XML, encoded with Unicode Transformation Format–8-bit (UTF-8), and relevant data sets to be standardized for efficient management and assessment. To support organizations in preparing claim data, the MoH has delivered data formatting guides and continuously updated them. The latest release is found in decision 4210/QD-BYT year 2017 (Requirements for standard and format of output data used in management, investigation and reimbursement of insurance-paid health care expenses [43]). Claims for groups of health services are recorded into separate data sets. Data sets from the same patient are linked through a unique ID number coded as *MA_LK*. The instructions include data item, data type, the allowed maximum length, and guidance for data collection.

The scope of the document is as follows:

- Guidance for claims of general health services
- Guidance for claims of medications
- Guidance for claims of medical procedures and supplies
- Guidance for claims of imaging diagnosis and laboratory test services
- Guidance for claims of follow-up interventions
- The official groups of services by cost
- The official hospital department codes
- The official accident and injury codes
- The official medicine bid packages and package codes

The Terminology and Service Coding System Version 6

The Service Coding System for Health Care Management and Health Insurance Reimbursement (decision 7603/QD-BYT year 2018 on the Service Coding System for Health Care Management and Health Insurance Reimbursement version 6 [48]), or the terminology and service coding system, was first piloted in 3 cities in Vietnam in 2015 [60]. Since then, it has been continuously updated, with version 6 [48] being the most up-to-date. The system features nomenclatures of health services, their standardized codes, regulated prices, and the ICD-10 codes with Vietnamese instruction. The terminology and service coding system has been made compulsory for use in health information systems of state health facilities and web-based health insurance claim activities. The 11 nomenclature groups provided in version 6 are as follows:

- Technical service codes
- Department codes
- Inpatient bed codes
- Half-day bed codes in chemotherapy and radiotherapy
- Medication codes
- Traditional medication codes
- Traditional diagnosis codes
- Medical supply codes
- Blood product codes
- ICD-10 codes
- Laboratory test codes

Cybersecurity in Health Organizations: The Cybersecurity Guidance

In October 2014, the MoH promulgated decision 4159/QD-BYT year 2014 (Guidance on ensuring the security of electronic health data in health organizations [37]) detailing the required measures to safeguard cybersecurity in health facilities. These regulations are applicable for organizations that conduct health IT activities in the web-based environment, including managing, using, storing, and exchanging health information. In general, the measures and recommendations addressed by this decision aim to maintain 3 characteristics (Textbox 6) in digital health implementation The 16 cybersecurity areas covered in the guidance can be found in Multimedia Appendix 6.

Textbox 6. Characteristics in digital health implementation safeguarded by The Cybersecurity Guidance.

Security

- Only authorized users can access the health information.
- Passwords and access keys are encrypted during access and transmission and are saved at the health care facilities.

Integrity

- Information can only be deleted or edited by authorized users. Information is preserved during storing and transmission.
- Integrity is maintained in the management, use, storage, and transmission of information, in which changes are not allowed without authorization from the administration unit.
- Measures to ensure integrity must be applied in accessing, entering, storing, using, processing, transferring, extracting, and recovering data.

Availability

- Uninterrupted operation of the information technology (IT) system is ensured.
- Information can be accessed quickly in response to authorized requests.
- Human resource personnel operating the IT system is ensured.
- Policies for managing and implementing the IT system are developed, publicized, and adhered to (Multimedia Appendix 6).

LISs: The LIS Guidance

Decision 3725/QD-BYT year 2017 (guideline for implementation of LIS in health care facilities [44]) guides the technical features and functionalities of an LIS used in health care facilities.

The guideline first highlights the essential points that organizations need to consider when implementing or developing an LIS:

- LIS's capability to communicate and integrate with other information systems in the hospital or health care system.
- Laboratory test results and related information can be exchanged and linked between different laboratories in a facility and with other facilities.
- LIS's design and technical documents to allow for easy repair, maintenance, and upgrade.
- Technical standards recognized in circular 22/2013/BTTTT and decision 2035/QD-BYT year 2013, where applicable, must be adopted for LISs.
- Advanced technologies are encouraged for implementing LISs.
- Facilities must ensure the necessary conditions for managing and operating LISs such as safety, data security, IT infrastructure, and human resources (Multimedia Appendix 7).

HIMS: The HIMS Guidance

In 2006, the MoH released the decision 5573/QD-BYT year 2006 (Requirements and functional modules for hospital management software (HMS) [35]) to guide the adoption of HMS in Vietnam's public and private health care facilities. The guideline addresses the key requirements that an HMS should meet in the regulatory and clinical context of Vietnam. An HMS's functional modules and detailed instructions for each module are also presented in the document (Multimedia Appendix 8).

Electronic Health Records

Overview

The EHR system aimed to create and maintain a lifetime EHR for every Vietnamese resident. It differed from EMR systems that are managed by hospitals. EHR contents were developed based on the paper-based personal health record's standardized template. Each EHR profile is identified by a unique health ID number.

The EHR Scheme

Overview

In decision 5349/QD-BYT year 2019 (The EHR Scheme [51]), the MoH presents the scheme for setting up and implementing the EHR system on a national scale. Accordingly, an EHR is the digital version of health records that are created, displayed, updated, stored, and exchanged using electronic devices. Health records are the medical documents that keep health care information of a person for their lifetime and are regulated by the MoH. The ultimate goal of this scheme is to provide every Vietnamese citizen with an EHR while step-by-step establishing a population health database in the National Health Data Center.

The scheme requires the manufacturers to apply specific standards for the EHR software. Textbox 7 lists the requirements for EHR systems according to the EHR Scheme.

Every EHR profile will have a national health ID number that is unique for every citizen. The National Health ID system is guided by decision 2153/QD-BYT year 2020 [54].

For data ownership, the government's administrative bodies (the MoH and the relevant Provincial Department of Health) own and manage the data generated during EHR use. EHR providers or EHR developers have responsibilities to hand over data, software source codes, and other tools that allow the MoH or the Department of Health to continue the EHR service with another provider.

Textbox 7. Summary of requirements for electronic health record systems.

Category and requirements

- The electronic health record (EHR) software's design must meet the required standards:
  - The EHR software can collect all the data in a standard paper health record guided in decision 831/QD-BYT year 2017 (the standard template of personal health record for primary health care)
  - The EHR software can export data to XML files that follow the guidance set out in decision 4210/QD-BYT year 2017 (Standards for claim data submitted to the web-based health insurance system)
  - The EHR software can satisfy the regulations in circular 48/2017/TT-BYT year 2017 (regulations for data sharing in health insurance claim)
  - Compatible with Health Level 7 standards
  - The EHR software can use the National Health ID system and is interoperable with related health information systems
- EHR functional modules
  - Health service provision functionality
  - Administrative management functionality
  - Information infrastructure management functionality
- Personal information protection
  - Protection of personal data in EHRs abides by Section 2, Chapter II of the Electronic Data Safety Law

Per the timeline announced in the scheme, the MoH aims to reach an EHR coverage of 80% in the population of central government provinces by the end of 2020 and a coverage of 95% in Vietnam's population by the end of 2025. Following are the steps to install and implement EHRs:

- Developing EHR software
- Installing EHR software in health facilities
- Providing EHR training for health workers
- Creating EHRs and collecting data for EHRs using the pre-existing data at health facilities or via interviews
- Continuously updating data to EHRs at health facilities
- Making use of data collected in EHRs
- Maintaining the EHR system

Technical documents and guidance to operate and manage EHRs will be built to guide EHR use at health facilities, including the following:

- Interoperability specifications with the National Health ID system and related health information systems.
- Policies for using, managing, operating, and ensuring data safety for the EHR system.
- Policies for creating, updating, and making use of EHRs.
- Financial mechanisms for maintaining and operating the EHR system.

The Health ID Regulation

Decision 2153/QD-BYT year 2020 (Regulations on creation, use, and management of health identification [54]), released in May 2020, announced that the social insurance ID will be the national health ID. These ID numbers are nationally unique for each patient and represent a set of identification data for that patient. Each identification data set includes the patient's full name, date of birth, gender, place of birth ID, and health insurance ID. Health ID is intended to be used with EHR, EMR, and other health information systems as a common patient identifier system.

Discussion

Principal Findings

This review explored the current state of digital health research and policies in Vietnam to inform the implementation of digital systems used in hospital care. Nearly half of the hospital-based studies were case reports of engineering solutions; 1 study assessed physicians' performance with assistance from a CDS software; 2 explored readiness to adopt EHRs; and 2 provided a high-level review of Vietnamese eHealth. The data analyzed in these studies were mostly collected in 2016 or earlier. These findings suggest a lack of research studies that investigate and inform the implementation of digital health systems in Vietnamese hospital settings.

Government policies in Vietnam have paid significant attention to digital health over the past 5 years. The MoH has set out specific regulations for EMR use and a maturity framework that shapes the development of digital health systems in hospitals.

Other national projects have occurred, such as the electronic insurance claim system, the nationwide rollout of an EHR, and a national health ID system. International standards such as HL7v2 messaging, ICD-10, and recently HL7's Fast Healthcare Interoperability Resources have been consistently recommended. Guidance documents have also addressed measures to ensure cybersecurity in health organizations. These policies are guided by an architectural framework and digital health strategies that ultimately aim to leverage IT, especially smart technology, in all aspects of the sector.

Of all the policies reviewed, the *Regulations for EMRs* and the *HIT Maturity Model* can be seen as the key policies for HIT application in Vietnam's hospitals. Although the previous national efforts focused on administrative systems such as the HIMS [35], these policies aimed to promote the adoption of clinical systems such as EMR, laboratory information management system, and RIS-PACS. At this nascent stage of HIT adoption, Vietnam has been focusing on building the necessary infrastructure and emphasizing the role of HIT in achieving a more efficient paperless environment. This is reflected in the recent digital health agenda, which highlighted the crucial role of EMRs in replacing paper medical records [49,53]. A major theme that the Regulations for EMRs addressed is the eligibility of transitioning to a completely paperless environment for EMR systems, as well as LIS and RIS-PACS [50]. Although this is a practical and well-defined motivation for EMR adoption, the World Health Organization recommended the true benefits of EMRs, such as improved data quality, timely access to information, and increased care quality, should be proven to encourage buy-in from health care workers [61]. Thus, future policies may seek to link EMR and HIT adoption with such benefits through clearly defined quality indicators to ensure meaningful HIT use. Lessons learned from some countries also recommend that financial incentives, when possible, should be made to increase EMR adoption [62-64].

The *HIT Maturity Model* defined a wide spectrum of HIT implementation levels that are applicable for hospitals of various digitization stages. Although the *Smart HIT Scheme* and the *Digital Transformation Scheme* aimed to establish smart hospitals (HIT level 6) in the next few years, there may be a need for further policy development that takes into account the Vietnamese context as an LMIC. The *Assessment of HIT Maturity* conducted in 2019 showed a large dispersion of HIT maturity across the top-level hospitals directly governed by the MoH, among which 19% (8/42) were only at HIT level 1, most hospitals (20/42, 48%) were at HIT level 3, and only 2% (1/42) of hospitals was at HIT level 5 [65]. A readiness assessment conducted by the MoH in 2017 revealed that among 36 MoH-governed hospitals, only 11 (31%) had EMR systems, 14 (39%) had some forms of CDS software, and over half of these hospitals lacked specialized IT workforce and information security capacity [66]. Given the low HIT maturity levels that these high-rank hospitals held, HIT implementation and readiness in smaller hospitals can be even poorer. It is likely that rural hospitals will struggle to reach even the most basic levels of digital maturity owing to a lack of funding, IT workforce, and infrastructure. Therefore, future policies may need to recommend further support for small and rural hospitals to reduce these gaps in digital maturity.

The 2 digital health strategies aim to promote AI in health care through applications such as CDSSs integrated with EHR systems, imaging diagnosis, and surgical support. New regulations, guidance, and standards are to be released to inform these adoptions [53]. Some of these intended policies are regulations for trials of novel digital health products, guidance for digital health technologies, and guidance for protecting web-based health data, to name a few. New regulation frameworks, along with an increased use of EMRs and EHRs in Vietnam's hospitals, will be important facilitators for the development and implementation of AI for hospital care in Vietnam. However, there is currently a lack of research to inform appropriate and sustainable use of these technologies in Vietnam. Local context factors such as disease burden, health care practices, and infrastructure constraints should be considered for future AI research and implementation, as highlighted by Schwalbe and Wahl [5], Wahl et al [6], and Alami et al [67]. Future digital health research should seek to investigate these factors to best guide HIT and AI development and application in Vietnam's hospitals.

Limitations

This review was limited to the context of hospital-based digital health systems and does not cover the more extensive academic literature on community-based systems and the use of national systems for monitoring infectious disease outbreaks and other public health interventions. Academic studies were only searched on international databases, hence this study did not cover publications in Vietnamese journals that are not registered in these databases.

At the time this study was conducted, the digital health policy landscape in Vietnam was experiencing significant changes with new policies likely to be released and the existing policies being subject to modifications over the next year. Therefore, this scoping review should be viewed as a snapshot of this area at the end of 2020.

Conclusions

Quantity and areas of research studies about digital health systems in hospitals in Vietnam are limited with little evidence to inform the implementation of new technologies in hospital care. The policies developed over the past 5 years to inform the adoption of digital health systems in Vietnam are comprehensive and will be useful for hospitals. They focus on guiding the basic functionalities and largely follow international standards and guidelines. Further research is needed to ensure that policies and guidelines can deal in detail with issues not encountered in HIC settings or that may be specific to the Vietnamese context.

Acknowledgments

This study was supported by the Wellcome Trust United Kingdom WT217650/Z/19/Z. Documents' full-texts on *Th Vi n Pháp Lu t* database was accessed through a registered paid account provided by the Oxford University Clinical Research Unit. Original government documents in Vietnamese were interpreted and summarized in English by DMT. English translations of some government documents purchased from *Th Vi n Pháp Lu t* were consulted during the interpretation. Search on Web of Science was conducted through the access from the Bodleian Health Care Libraries.

The members of the *Vietnam ICU Translational Applications Laboratory (VITAL)* group are as follows:

Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam: An Phuoc Luu, Chanh Quang Ho, Duc Hong Du, Duc Minh Tran, Dung Thi Phuong Nguyen, Giang Thi Nguyen, Hai Bich Ho, Hien Van Ho, Hung Manh Trinh, Huy Quang Nguyen, Khanh Nguyen Quoc Phan, Khoa Dinh Van Le, Kien Trung Dang, Lam Khanh Phung, Lieu Thi Pham, Ngoc Thanh Nguyen, Nhat Tran Huy Phung, Phuong Thanh Le, Quyen Than Ha Nguyen, Thanh Thi Le Nguyen, Thy Bui Xuan Doan, Trieu Trung Huynh, Trinh Huu Khanh Dong, Van Minh Tu Hoang, Van Thi Thanh Ninh, Vuong Lam Nguyen, Yen Minh Lam, Sayem Ahmed, Joseph Donovan, Ronald Geskus, Evelyne Kestelyn, Angela Mcbride, Guy Thwaites, Louise Thwaites, Hugo Turner, Jennifer Ilo Van Nuil, and Sophie Yacoub.

Hospital for Tropical Diseases, Ho Chi Minh City, Vietnam: Tam Thi Cao, Thuy Bich Duong, Duong Thi Hai Ha, Nghia Dang Trung Ha, Chau Buu Le, Thu Ngoc Minh Le, Thao Thi Mai Le, Tai Thi Hue Luong, Phu Hoan Nguyen, Viet Quoc Nguyen, Nguyen Thanh Nguyen, Phong Thanh Nguyen, Anh Thi Kim Nguyen, Hao Van Nguyen, Duoc Van Thanh Nguyen, Chau Van Vinh Nguyen, Oanh Kieu Nguyet Pham, Van Thi Hong Phan, Qui Tu Phan, Tho Vinh Phan, and Thao Thi Phuong Truong.

University of Oxford, Oxford, United Kingdom: David Clifton, Mike English, Shadi Ghiasi, Heloise Greeff, Jannis Hagenah, Ping Lu, Jacob McKnight, Chris Paton, and Tingting Zhu.

Imperial College London, London, United Kingdom: Pantellis Georgiou, Bernard Hernandez Perez, Kerri Hill-Cawthorne, Alison Holmes, Stefan Karolcik, Damien Ming, Nicolas Moser, and Jesus Rodriguez Manzano.

King's College London, London, United Kingdom: Alberto Gomez, Hamideh Kerdegari, Marc Modat, and Reza Razavi.

ETH Zurich, Zurich, Switzerland: Abhilash Guru Dutt, Walter Karlen, Michaela Verling, and Elias Wicki.

The University of Melbourne, Melbourne, Australia: Linda Denehy and Thomas Rollinson.

Authors' Contributions

All authors contributed to the design of the study, interpretation of the data, and revision of the manuscript before it is published. CLT, JM, and CP acquired the funding. DMT collected and analyzed the data, and wrote the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Literature search strategies.
[DOC File , 29 KB - jmir_v24i2e32392_app1.doc]

Multimedia Appendix 2
Circular 46/2018/TT-BYT on regulations for electronic medical records.
[DOC File , 46 KB - jmir_v24i2e32392_app2.doc]

Multimedia Appendix 3
Circular 53/2014/TT-BYT on required conditions for provision of health information technology activities.
[DOC File , 47 KB - jmir_v24i2e32392_app3.doc]

Multimedia Appendix 4
Circular 39/2017/TT-BTTTT on technical standards for information technology implementation in state organizations.
[DOC File , 169 KB - jmir_v24i2e32392_app4.doc]

Multimedia Appendix 5
Circular 48/2017/TT-BYT on regulations on data exchange in management and reimbursement of health insurance claims.
[DOC File , 37 KB - jmir_v24i2e32392_app5.doc]

Multimedia Appendix 6
Decision 4159/QD-BYT year 2014 on guidance on ensuring security of electronic health data in health organizations.
[DOC File , 72 KB - jmir_v24i2e32392_app6.doc]

Multimedia Appendix 7
Decision 3725/QD-BYT year 2017 on guideline for implementation of laboratory information system in health care facilities.
[DOC File , 59 KB - jmir_v24i2e32392_app7.doc]

Multimedia Appendix 8
Decision 5573/QD-BYT year 2006 on requirements and functional modules for hospital management software.
[DOC File , 121 KB - jmir_v24i2e32392_app8.doc]

References

1. Halamka JD, Tripathi M. The HITECH era in retrospect. N Engl J Med 2017 Sep 07;377(10):907-909. [doi: 10.1056/NEJMp1709851] [Medline: 28877012]
2. Price C, Green W, Suhomlinova O. Twenty-five years of national health IT: exploring strategy, structure, and systems in the English NHS. J Am Med Inform Assoc 2019 Mar 01;26(3):188-197 [FREE Full text] [doi: 10.1093/jamia/ocy162] [Medline: 30597001]
3. Friedman C, Rubin J, Brown J, Buntin M, Corn M, Etheredge L, et al. Toward a science of learning systems: a research agenda for the high-functioning Learning Health System. J Am Med Inform Assoc 2015 Jan;22(1):43-50 [FREE Full text] [doi: 10.1136/amiajnl-2014-002977] [Medline: 25342177]
4. Paton C, Kobayashi S. An open science approach to artificial intelligence in healthcare. Yearb Med Inform 2019 Aug 25;28(1):47-51 [FREE Full text] [doi: 10.1055/s-0039-1677898] [Medline: 31022753]
5. Schwalbe N, Wahl B. Artificial intelligence and the future of global health. Lancet 2020 May;395(10236):1579-1586. [doi: 10.1016/s0140-6736(20)30226-9]
6. Wahl B, Cossy-Gantner A, Germann S, Schwalbe NR. Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings? BMJ Glob Health 2018 Aug 29;3(4):e000798. [doi: 10.1136/bmjgh-2018-000798] [Medline: 30233828]
7. Muinga N, Magare S, Monda J, English M, Fraser H, Powell J, et al. Digital health Systems in Kenyan Public Hospitals: a mixed-methods survey. BMC Med Inform Decis Mak 2020 Jan 06;20(1):2 [FREE Full text] [doi: 10.1186/s12911-019-1005-7] [Medline: 31906932]
8. Muinga N, Magare S, Monda J, Kamau O, Houston S, Fraser H, et al. Implementing an open source electronic health record system in Kenyan health care facilities: case study. JMIR Med Inform 2018 Apr 18;6(2):e22 [FREE Full text] [doi: 10.2196/medinform.8403] [Medline: 29669709]
9. McCool J, Dobson R, Muinga N, Paton C, Pagliari C, Agawal S, et al. Factors influencing the sustainability of digital health interventions in low-resource settings: lessons from five countries. J Glob Health 2020 Dec;10(2):020396 [FREE Full text] [doi: 10.7189/jogh.10.020396] [Medline: 33274059]
10. Paton C, Karopka T. The role of free/libre and open source software in learning health systems. Yearb Med Inform 2017 Aug 19;26(01):53-58. [doi: 10.1055/s-0037-1606527]
11. Kumar P, Paton C, Kirigia D. I've got 99 problems but a phone ain't one: electronic and mobile health in low and middle income countries. Arch Dis Child 2016 Dec;101(10):974-979. [doi: 10.1136/archdischild-2015-308556] [Medline: 27296441]
12. Manya A, Braa J, Øverland L, Titlestad O, Mumo J, Nzioka C. National roll out of District Health Information Software (DHIS 2) in Kenya, 2011--central server and cloud based infrastructure. In: Proceedings of the IST-Africa 2012 Conference. 2012 Presented at: IST-Africa 2012 Conference; May 09 - 11, 2012; Dar es Salaam, Tanzania URL: https://tinyurl.com/2vtd422p [doi: 10.1109/istafrica.2016.7530690]
13. Index - FHIR v4. Health Level Seven International. URL: https://www.hl7.org/fhir/ [accessed 2021-12-05]
14. Mandel JC, Kreda DA, Mandl KD, Kohane IS, Ramoni RB. SMART on FHIR: a standards-based, interoperable apps platform for electronic health records. J Am Med Inform Assoc 2016 Dec;23(5):899-908 [FREE Full text] [doi: 10.1093/jamia/ocv189] [Medline: 26911829]
15. Dinh MN, Nygate J, Tu VH, Thwaites CL. New technologies to improve healthcare in low- and middle-income countries: Global Grand Challenges satellite event, Oxford University Clinical Research Unit, Ho Chi Minh City, 17th-18th September 2019. Wellcome Open Res 2020 Jun 15;5:142. [doi: 10.12688/wellcomeopenres.16008.1]
16. Turner HC, Hao NV, Yacoub S, Hoang VM, Clifton DA, Thwaites GE, et al. Achieving affordable critical care in low-income and middle-income countries. BMJ Glob Health 2019;4(3):e001675 [FREE Full text] [doi: 10.1136/bmjgh-2019-001675] [Medline: 31297248]
17. The Sixth Conference of the 12th Central Executive Committee on enhancing the protection, care and enhanceding the people's health in the new situation. The Central Committee, Communist Party of Vietnam. 2017. URL: https://tinyurl.com/ycxccrcw [accessed 2021-12-05]

18. Koh D. Vietnam to deploy EHR nationwide in July. Healthcare IT News. 2019. URL: https://www.healthcareitnews.com/news/asia-pacific/vietnam-deploy-ehr-nationwide-july [accessed 2021-12-05]
19. Digital healthcare to create motivation for investment. Vietnam Investment Review. 2020. URL: https://www.vir.com.vn/digital-healthcare-to-create-motivation-for-investment-81881.html [accessed 2021-12-05]
20. Anderson S, Allen P, Peckham S, Goodwin N. Asking the right questions: scoping studies in the commissioning of research on the organisation and delivery of health services. Health Res Policy Syst 2008 Jul 09;6:7 [FREE Full text] [doi: 10.1186/1478-4505-6-7] [Medline: 18613961]
21. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
22. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
23. Thu Vien Pháp Luat. thuvienphapluat.vn. URL: https://thuvienphapluat.vn/ [accessed 2021-12-06]
24. Vu D, Nguyen D. Design of laboratory information system for healthcare in Vietnam BK-LIS. In: Proceedings of the International Conference on Communications and Electronics 2010. 2010 Presented at: International Conference on Communications and Electronics 2010; Aug. 11-13, 2010; Nha Trang, Vietnam p. 110-114. [doi: 10.1109/icce.2010.5670692]
25. Vu H, Nguyen T, Pham N, Hoang H, Pham TV. A design of renal dataflow control and patient record management system for renal department environment in Vietnam. In: Proceedings of the The Third International Conference on the Development of Biomedical Engineering in Vietnam. 2010 Presented at: The Third International Conference on the Development of Biomedical Engineering in Vietnam; January 11–14, 2010; Ho Chi Minh City, Vietnam p. 98-101. [doi: 10.1007/978-3-642-12020-6_24]
26. Hai V, Thuan N, Ngoc P. Automatic retrieving data from medical equipment to create Electronic Medical Records (EMRs) for an e-hospital model in Vietnam. In: Proceedings of the 5th International Symposium on Bio- and Medical Informatics and Cybernetics. 2011 Presented at: 5th International Symposium on Bio- and Medical Informatics and Cybernetics; Jul 19-22, 2011; Orlando, FL, USA p. 152-157 URL: https://www.iiis.org/cds2011/cd2011sci/bmic_2011/paperspdf/ba989ps.pdf
27. Vo C, Ho B, Nguyen H. Towards VNUMED for healthcare research activities in Vietnam. ERCIM News. 2019. URL: https://ercim-news.ercim.eu/en118/special/towards-vnumed-for-healthcare-research-activities-in-vietnam [accessed 2021-12-05]
28. VO C, Cao T, Doan N, Tu S, Nguyen A, Nguyen B, et al. EMRVisualization for patient progress tracking: interface design and system implementation. Int J Affect Eng 2019;18(3):117-128. [doi: 10.5057/ijae.ijae-d-18-00025]
29. Nguyen Q, Naguib R, Tawfik H, Phuong N. Strategic challenges facing user- and patient-centred e-health in Vietnam. Int J Health Technol Manag 2012;13(5/6):378. [doi: 10.1504/ijhtm.2012.052551]
30. Lin M, Brooks T, Miller A, Sharp J, Hai L, Nguyen T, et al. English-based pediatric emergency medicine software improves physician test performance on common pediatric emergencies: a multicenter study in Vietnam. West J Emerg Med 2013 Sep;14(5):471-476 [FREE Full text] [doi: 10.5811/westjem.2013.4.15090] [Medline: 24106545]
31. Hochwarter S, Cuong DD, Chuc NT, Larsson M. Towards an electronic health record system in Vietnam: a core readiness assessment. J Health Inform Dev Countries 2014;8(2) [FREE Full text]
32. Landgraf K, Kakkar R, Meigs M, Jankauskas P, Phan T, Nguyen V, et al. Open-source LIMS in Vietnam: the path toward sustainability and host country ownership. Int J Med Inform 2016 Sep;93:92-102. [doi: 10.1016/j.ijmedinf.2016.06.010] [Medline: 27435952]
33. Nguyen KT. Electronic health record readiness assessment in Thai Binh Hospital, Vietnam. Int J Sci Res 2018;7(3):5 [FREE Full text]
34. Vuong Q, Ho M, Vuong T, La V, Ho M, Nghiem KP, et al. Artificial intelligence vs. natural stupidity: evaluating ai readiness for the vietnamese medical information system vien pháp luat. J Clin Med 2019 Feb 01;8(2):168 [FREE Full text] [doi: 10.3390/jcm8020168] [Medline: 30717268]
35. Quy t đ nh: V vi c ban hành "tiêu chí ph n m m và n i dung m t s phân h ph n m m tin h c qu n lý b nh vi n". Vietnam Ministry of Health. 2006. URL: https://thuvienphapluat.vn/van-ban/Cong-nghe-thong-tin/Quyet-dinh-5573-QD-BYT-Tieu-chi-phan-mem-noi-dung-phan-he-111827.aspx [accessed 2021-12-06]
36. Quy t đ nh : V vi c công b danh m c k thu t v ng d ng công ngh thông tin trong lĩnh v c y t . Vietnam Ministry of Health. 2013. URL: https://tinyurl.com/2p96dap9 [accessed 2021-12-06]
37. Quy t đ nh : Ban hành quy đ nh v đ m b o an toàn thông tin y t đi n t t i các đ n v trong ngành y t . Vietnam Ministry of Health. 2014. URL: https://tinyurl.com/yckmv8nv [accessed 2021-12-06]
38. Thông t quy đ nh đi u ki n ho t đ ng y t trên môi tr ng m ng. Vietnam Ministry of Health. 2014. URL: https://tinyurl.com/3wvaaymj [accessed 2021-12-06]
39. Quy t đ nh: Ban hành h ng d n xây d ng n i quy an toàn, an ninh thông tin trong các đ n v trong ngành y t . Vietnam Ministry of Health. 2015. URL: https://tinyurl.com/2kpysywa [accessed 2021-12-06]
40. Quy t đ nh : Ban hành quy trình ph n ng v i các s c an toàn, an ninh thông tin t i các đ n v trong ngành y t . Vietnam Ministry of Health. 2015. URL: https://tinyurl.com/5xd8dvfs [accessed 2021-12-06]

41. Quy t đ nh : V vi c phê duy t mô hình ki n trúc t ng th h th ng thông tin khám ch a b nh b o hi m y t . Vietnam Ministry of Health. 2016. URL: https://tinyurl.com/22vtx3bk [accessed 2021-12-06]
42. Quy t đ nh : Ban hành c ng ti p nh n d li u h th ng thông tin giám đ nh b o hi m y t phiên b n 2.0. Vietnam Social Security. 2016. URL: https://tinyurl.com/ms9yp3xx [accessed 2021-12-06]
43. Quy t đ nh : V vi c quy đ nh chu n và đ nh d ng d li u đ u ra s d ng trong qu n lý, giám đ nh và thanh toán chi phí khám b nh, ch a b nh b o hi m y t . Vietnam Ministry of Health. 2017. URL: https://tinyurl.com/yckjc4az [accessed 2021-12-06]
44. Quy t đ nh : Ban hành h ng d n tri n khai h th ng thông tin qu n lý xét nghi m t i các c s khám b nh, ch a b nh. Vietnam Ministry of Health. 2017. URL: https://tinyurl.com/ycksffmd [accessed 2021-12-06]
45. Thông t : Ban hành b tiêu chí ng d ng công ngh thông tin t i các c s khám b nh, ch a b nh. Vietnam Ministry of Health. 2017. URL: https://tinyurl.com/kctm4kf4 [accessed 2021-12-06]
46. Thông t : Ban hành danh m c tiêu chu n k thu t v ng d ng công ngh thông tin trong c quan nhà n c. Vietnam Ministry of Information and Communication. 2017. URL: https://tinyurl.com/2p9xrkmd [accessed 2021-12-06]
47. Thông t : Quy đ nh trích chuy n d li u đi n t trong qu n lý và thanh toán chi phí khám b nh, ch a b nh b o hi m y t . Vietnam Ministry of Health. 2017. URL: https://tinyurl.com/msvu3wfj [accessed 2021-12-06]
48. Quy t đ nh : V vi c ban hành b mã danh m c dùng chung áp d ng trong qu n lý khám b nh, ch a b nh và thanh toán b o hi m y t (phiên b n s 6). Vietnam Ministry of Health. 2018. URL: https://tinyurl.com/2p94furr [accessed 2021-12-06]
49. Quy t đ nh : Phê duy t đ án ng d ng và phát tri n công ngh thông tin y t thông minh giai đo n 2019 -2025. Vietnam Ministry of Health. 2019. URL: https://tinyurl.com/2p885fhk [accessed 2021-12-06]
50. Thông t : Quy đ nh h s b nh án đi n t . Vietnam Ministry of Health. 2018. URL: https://tinyurl.com/2m5rj3tb [accessed 2021-12-06]
51. Quy t đ nh : Phê duy t k ho ch tri n khai h s s c kh e đi n t . Vietnam Ministry of Health. 2019. URL: https://tinyurl.com/bddh3dwb [accessed 2021-12-06]
52. Quy t đ nh : Ban hành ki n trúc chính ph đi n t b y t phiên b n 2.0. Vietnam Ministry of Health. 2019. URL: https://tinyurl.com/2p8ha5mt [accessed 2021-12-06]
53. Quy t đ nh : Phê duy t ch ng trình chuy n đ i s y t đ n năm 2025, đ nh h ng đ n năm 2030. Vietnam Ministry of Health. 2020. URL: https://tinyurl.com/mrnpz8m5 [accessed 2021-12-06]
54. Quy t đ nh : Ban hành quy ch xác l p, s d ng và qu n lý mã đ nh danh y t . Vietnam Ministry of Health. 2020. URL: https://tinyurl.com/4xty4zy4 [accessed 2021-12-06]
55. Electronic Medical Record Adoption Model (EMRAM). Healthcare Information and Management Systems Society. 2018. URL: https://www.himssanalytics.org/emram [accessed 2021-12-06]
56. Quy t đ nh : V vi c ban hành m u h s qu n lý s c kh e cá nhân ph c v chăm sóc s c kh e ban đ u. Vietnam Ministry of Health. 2017. URL: https://tinyurl.com/53an76d7 [accessed 2021-12-06]
57. Tên văn b n: Ki n trúc Chính ph đi n t B Y t phiên b n 1.0. Vietnam Ministry of Health. 2015. URL: https://ehealth.gov.vn/Index.aspx?action=Detail&MenuChildID=402&Id=4235 [accessed 2021-12-06]
58. Quy t đ nh : C p nh t ki n trúc chính ph đi n t b y t phiên b n 1.0. Vietnam Ministry of Health. 2018. URL: https://tinyurl.com/3a7x8s8a [accessed 2021-12-06]
59. Lu t : An toàn thông tin m ng. The National Assembly. 2015. URL: https://thuvienphapluat.vn/van-ban/Cong-nghe-thong-tin/Luat-an-toan-thong-tin-mang-2015-298365.aspx [accessed 2021-12-06]
60. QUy t đ nh : V vi c ban hành quy đ nh t m th i b mã danh m c dùng chung đ thí đi m áp d ng trong khám b nh, ch a b nh và thanh toán b o hi m y t t i t nh thái nguyên, b c ninh và thành ph h i phòng. Vietnam Ministry of Health. 2015. URL: https://thuvienphapluat.vn/van-ban/Bao-hiem/Quyet-dinh-2182-QD-BYT-bo-ma-danh-muc-dung-chung-kham-chua-benh-thanh-toan-bao-hiem-y-te-2015-294286.aspx [accessed 2021-12-06]
61. World Health Organization. Electronic Health Records: Manual for Developing Countries. Manila: WHO Regional Office for the Western Pacific; 2006:1-76.
62. Liang J, Li Y, Zhang Z, Shen D, Xu J, Zheng X, et al. Adoption of Electronic Health Records (EHRs) in China during the past 10 years: consecutive survey data analysis and comparison of sino-american challenges and experiences. J Med Internet Res 2021 Feb 18;23(2):e24813 [FREE Full text] [doi: 10.2196/24813] [Medline: 33599615]
63. Kruse CS, Kristof C, Jones B, Mitchell E, Martinez A. Barriers to electronic health record adoption: a systematic literature review. J Med Syst 2016 Dec;40(12):252 [FREE Full text] [doi: 10.1007/s10916-016-0628-9] [Medline: 27714560]
64. Kim Y, Jung K, Park Y, Shin D, Cho S, Yoon D, et al. Rate of electronic health record adoption in South Korea: a nation-wide survey. Int J Med Inform 2017 May;101:100-107 [FREE Full text] [doi: 10.1016/j.ijmedinf.2017.02.009] [Medline: 28347440]
65. Công b m c đ ng d ng công ngh thông tin trong các b nh vi n tr c thu c b y t năm 2019 theo các tiêu chí c a thông t 54/2017/tt-byt. Vietnam Ministry of Health. 2020. URL: https://ehealth.gov.vn/Index.aspx?action=News&newsId=49934 [accessed 2021-12-06]
66. V vi c phê duy t và công b báo cáo m c đ s n sàng cho phát tri n và ng d ng công ngh thông tin ngành y t năm 2017. Vietnam Ministry of Health. 2017. URL: https://tinyurl.com/2p9e6cba [accessed 2021-12-06]

67. Alami H, Rivard L, Lehoux P, Hoffman SJ, Cadeddu SB, Savoldelli M, et al. Artificial intelligence in health care: laying the foundation for responsible, sustainable, and inclusive innovation in low- and middle-income countries. Global Health 2020 Jun 24;16(1):52 [FREE Full text] [doi: 10.1186/s12992-020-00584-1] [Medline: 32580741]

Abbreviations

AI: artificial intelligence
CDR: clinical data repository
CDS: clinical decision support
CDSS: clinical decision support system
EHAF: eHealth architectural framework
EHR: electronic health record
EMR: electronic medical record
HIC: high-income country
HIMS: Hospital Information Management System
HIS: hospital information system
HIT: health information technology
HL7: Health Level 7
HMS: hospital management software
ICD-10: International Classification of Diseases, Tenth Revision
IT: information technology
LIS: laboratory information system
LMIC: low- and middle-income country
MIC: Ministry of Information and Communications
MoH: Ministry of Health
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
RIS-PACS: radiology information system and picture archiving and communication systems
SDMX-HD: Statistical Data and Metadata Exchange–Health Domain
UTF-8: Unicode Transformation Format–8-bit
VSS: Vietnam Social Security

Edited by R Kukafka; submitted 29.07.21; peer-reviewed by T Wang, R Dekova; comments to author 31.08.21; revised version received 23.09.21; accepted 21.10.21; published 09.02.22.

Please cite as:
Tran DM, Thwaites CL, Van Nuil JI, McKnight J, Luu AP, Paton C, Vietnam ICU Translational Applications Laboratory (VITAL)
Digital Health Policy and Programs for Hospital Care in Vietnam: Scoping Review
J Med Internet Res 2022;24(2):e32392
URL: https://www.jmir.org/2022/2/e32392
doi:10.2196/32392
PMID:35138264

Review

The Effectiveness of Virtual Reality Exposure–Based Cognitive Behavioral Therapy for Severe Anxiety Disorders, Obsessive-Compulsive Disorder, and Posttraumatic Stress Disorder: Meta-analysis

Inge van Loenen[1,2*], MSc; Willemijn Scholten[1,2*], PhD; Anna Muntingh[1,2], PhD; Johannes Smit[1,2], PhD; Neeltje Batelaan[1,2], MD, PhD

[1]GGZ inGeest Specialized Mental Health Care, Amsterdam, Netherlands

[2]Amsterdam UMC, Vrije Universiteit Amsterdam, Psychiatry, Amsterdam Public Health Research Institute, Amsterdam, Netherlands

[*]these authors contributed equally

Corresponding Author:
Willemijn Scholten, PhD
GGZ inGeest Specialized Mental Health Care
Oldenaller 1
Amsterdam, 1081HJ
Netherlands
Phone: 31 207884666
Fax: 31 207884677
Email: w.scholten@ggzingeest.nl

Abstract

Background: In recent years, virtual reality exposure–based cognitive behavioral therapy (VRE-CBT) has shown good treatment results in (subclinical) anxiety disorders and seems to be a good alternative to exposure in vivo in regular cognitive behavioral therapy (CBT). However, previous meta-analyses on the efficacy of VRE-CBT on anxiety disorders have included studies on specific phobias and subthreshold anxiety; therefore, these results may not be generalizable to patients with more severe and disabling anxiety disorders.

Objective: The objective of our study is to determine the efficacy of VRE-CBT on more severe anxiety disorders, excluding specific phobias and subthreshold anxiety disorders. Meta-analyses will be conducted to examine the efficacy of VRE-CBT versus waitlist and regular CBT. Our secondary objectives are to examine whether the efficacy differs according to the type of anxiety disorder, type of recruitment, and type of VRE-CBT (virtual reality exposure either with or without regular CBT). Furthermore, attrition in VRE-CBT and CBT will be compared.

Methods: Studies published until August 20, 2020, were retrieved through systematic literature searches in PubMed, PsycINFO, and Embase. We calculated the effect sizes (Hedges *g*) for the difference between the conditions and their 95% CIs for posttest and follow-up measurements in a random effects model. A separate meta-analysis was performed to compare attrition between the VRE-CBT and CBT conditions.

Results: A total of 16 trials with 817 participants were included. We identified 10 comparisons between VRE-CBT and a waitlist condition and 13 comparisons between VRE-CBT and a CBT condition. With regard to risk of bias, information on random sequence generation, allocation concealment, and risk of bias for selective outcome reporting was often absent or unclear. The mean effect size of VRE-CBT compared with waitlist (n_{co}=10) was medium and significant, favoring VRE-CBT (Hedges *g*=−0.490, 95% CI −0.82 to −0.16; *P*=.003). The mean effect size of VRE-CBT compared with CBT (n_{co}=13) was small and nonsignificant, favoring CBT (Hedges *g*=0.083, 95% CI −0.13 to 0.30; *P*=.45). The dropout rates between VRE-CBT and CBT (n_{co}=10) showed no significant difference (odds ratio 0.79, 95% CI 0.49-1.27; *P*=.32). There were no indications of small study effects or publication bias.

Conclusions: The results of our study show that VRE-CBT is more effective than waitlist and as effective as CBT in the treatment of more severe anxiety disorders. Therefore, VRE-CBT may be considered a promising alternative to CBT for patients with more severe anxiety disorders. Higher-quality randomized controlled trials are needed to verify the robustness of these findings.

(J Med Internet Res 2022;24(2):e26736) doi:10.2196/26736

KEYWORDS

anxiety disorders; virtual reality; virtual reality exposure therapy; cognitive behavioral therapy; meta-analysis; mobile phone

Introduction

Background

Exposure-based cognitive behavioral therapy (CBT) is known as the *gold standard* for treatment of patients with anxiety disorders as its efficacy has been well-established by extensive research [1,2]. Despite its efficacy, potential disadvantages concern CBT being time- and cost-inefficient, impracticable, or considered too aversive by patients [3]. Virtual reality exposure–based cognitive behavioral therapy (VRE-CBT) offers an alternative to regular CBT and has several advantages. It allows immersion within a feared virtual environment that is tailored to the individual patient while being offered within a convenient and safe clinical setting. In VRE-CBT, the therapist can ensure complete control over the content and dose of feared stimuli and therefore optimize individualized pacing through exposures. Moreover, each step of VRE-CBT can be repeated as often as needed before proceeding to the next feared situation, thereby facilitating the complex processes responsible for the effect of exposure therapy. Attrition rates between VRE-CBT and CBT do not seem to differ [4]. However, it was shown that most of a clinical sample with phobic disorders (76%) preferred VRE-CBT over exposure in vivo, and refusal rates of those who were offered VRE-CBT (3%) were substantially lower compared with refusal rates of those who were offered exposure in vivo (27%) [5].

In recent years, many studies have been conducted on the efficacy of VRE-CBT on anxiety disorders and have shown good treatment results. A total of 5 meta-analyses on the efficacy of VRE-CBT versus control conditions, consisting mostly of studies on specific phobias, have been published. They consistently show a clear superiority of VRE-CBT versus nonactive control groups (ie, waitlist) [6-10] and similar [6,7,9,10] or larger [8] effects of VRE-CBT versus CBT conditions incorporating in vivo exposure. In addition, Opris et al [6] showed good stability of VRE-CBT results over time, similar to those of CBT, and a dose-response relationship with more sessions yielding a larger treatment effect. Although it has been questioned whether VRE-CBT induces significant behavior changes in real life [8], Morina et al [7] showed that treatment gains after VRE-CBT generalize well to real-life situations as measured by means of behavioral laboratory tests and recordings of behavioral activities in real life [7]. Considering the advantages of VRE-CBT and its promising results in previous meta-analyses, it is not surprising that a significant expansion of VRE-CBT in mental health care is both predicted and advocated to occur in the coming years [11]. However, these meta-analyses have limited generalizability to patients with more severe and disabling anxiety symptoms—a diagnosis of an anxiety disorder according to the criteria of the Diagnostic and Statistical Manual of Mental Disorders, Fourth and Fifth Edition (DSM-IV and DSM-V, respectively), and the International Classification of Diseases, 10th and 11th Revision (ICD-10 and ICD-11, respectively), established by a structured diagnostic interview, excluding specific phobias. First, most previous meta-analyses included studies that completely [7], largely, or partly [6,8-10] focused on specific phobias. However, specific phobias are generally less severe than more disabling anxiety disorders such as social anxiety disorder (SAD), posttraumatic stress disorder (PTSD), panic disorder (PD), and agoraphobia (AGO). Second, the inclusion of studies in previous meta-analyses was not restricted to samples with a formal anxiety disorder diagnosis (following DSM or ICD criteria), thereby including subthreshold anxiety disorders that are less severe. Carl et al [9] also stated that "The applications for clinical settings are clearer when conclusions can be drawn from clinical samples." The experience with and effects of VRE-CBT may be different for patients with more severe symptoms. On the one hand, they might be more reluctant to expose themselves to feared situations with regular CBT and may favor and benefit most from a controlled exposure therapy setting, such as in VRE-CBT. On the other hand, treating these more severe anxiety disorders with VRE-CBT may be far more challenging than treating milder symptoms. VRE-CBT for the treatment of more severe anxiety disorders requires more varied and elaborate virtual settings while still ensuring a high sense of presence (the feeling that one is actually present in the virtual world). For example, rather than fear of a specific situation or object, treatment of AGO generally requires exposure to various public places (eg, streets, shops, and public transport). In addition, for SAD, it may be essential that VRE-CBT also incorporates interaction through verbal and nonverbal behavior of others in response to the behavior of the patient as SAD centers on the perception of negative evaluation. In addition to the predominance of studies targeting specific phobias, most previous meta-analyses also included nonrandomized studies, thereby possibly leading to biased results. Ultimately, there is a lack of evidence for the efficacy of VRE-CBT in the treatment of more severe and formal anxiety disorders. Hence, a careful examination of research is necessary to prevent unjustified expansion of VRE-CBT in clinical settings.

We conducted a meta-analysis to determine whether the positive results of VRE-CBT as reported in earlier meta-analyses are sustained when focusing solely on more severe anxiety disorders (according to DSM-IV, DSM-V, ICD-10, or ICD-11 criteria): PD with or without AGO, AGO, SAD, generalized anxiety disorder (GAD), obsessive-compulsive disorder (OCD), or PTSD, limiting our inclusion criteria to studies with a formal diagnosis to avoid heterogeneity in our sample.

Objectives

Our main objective is to examine the efficacy of VRE-CBT on anxiety severity compared with (1) waitlist and (2) CBT at posttest and follow-up measurements (if available). Our secondary objectives are to examine whether efficacy differs according to the type of anxiety disorder, type of recruitment, and type of VRE-CBT (virtual reality exposure [VRE] with or

without regular CBT). Furthermore, the feasibility of VRE-CBT has been evaluated by comparing the attrition of VRE-CBT and CBT.

Methods

Study Selection

Studies published until August 20, 2020, were retrieved through systematic literature searches in PubMed, PsycINFO, and Embase (see Multimedia Appendix 1 for the complete search strings) for studies that randomized patients with anxiety disorders into VRE-CBT, active control condition, or an inactive control condition and compared the effects of these treatments. A librarian and IvL conducted the search by combining search terms indicative of anxiety disorders (PD with or without AGO, AGO, GAD, SAD, OCD, and PTSD) with terms indicative of VRE therapy. Furthermore, we checked the reference lists of the retrieved articles and previous meta-analyses and reviews for additional studies.

The inclusion criteria were (1) adult patients (aged >18 years); (2) at least one VRE-CBT condition; (3) random assignment to conditions; (4) comparison with waitlist or CBT without virtual reality (VR); (5) measure of outcome related to anxiety; (6) a primary diagnosis of an anxiety disorder according to DSM-IV, DSM-V, ICD-10, or ICD-11 criteria established by a structured diagnostic interview, excluding specific phobias; and (7) original empirical findings. Consistent with the DSM-IV or DSM-V classification, PTSD and OCD were also included even though these are no longer classified as anxiety disorders in the DSM-V. No date restrictions were applied. Articles consisting of only abstracts or not presenting original data were excluded, as were studies with a crossover design. No language restrictions were applied.

In total, 2 authors (WS and IvL) independently assessed the list of titles and abstracts that resulted from the literature search for eligibility. WS and IvL independently examined full texts and selected eligible randomized controlled trials (RCTs). Discrepancies were resolved by consulting and discussing with a third author (NB). This meta-analysis was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [12].

Data Extraction

In total, 2 authors (WS and IvL) independently extracted the data from the selected articles. This information included characteristics of the study population (mean age and range and gender), classification instrument, recruitment method, primary outcome, and conditions (sample size, number of dropouts, number of sessions, and follow-up in months for each condition).

For studies using more than one validated outcome measure, we selected the reported primary outcome measure. When undefined, we selected the domain-specific outcome that was most frequently used in the included trials for that type of anxiety disorder as the primary outcome measure.

For each study, the mean and SD of the primary outcome on posttest and follow-up measurements were used to calculate the effect sizes. For the study of Pitti et al [13], the effect size was calculated using the mean and *P* value of the between-group comparison as SDs were not available. Discrepancies in data extraction were resolved through discussion and by studying the original article. In case of a lack of clarity or missing data, the authors of selected studies were contacted.

Risk of Bias Assessment

In total, 2 authors (WS and IvL) independently assessed the risk of bias in the included studies. Differences were resolved through discussion. Five areas of risk of bias according to the Cochrane tool for assessing risk of bias were rated [14]: (1) random sequence generation, (2) allocation concealment, (3) blinding of outcome assessment, (4) incomplete outcome data, and (5) selective reporting. The criterion of blinding of participants and personnel was omitted as blinding is not possible in these psychological interventions. We assessed all areas as having low, high, or unclear risk of bias. Incomplete outcome data were assessed by evaluating attrition rates in all the comparison groups. Attrition was defined as all randomized patients who dropped out of the treatment. If attrition rates were >30% and no appropriate methods were used to minimize biased comparisons among groups (intention-to-treat analyses), the study was coded as being at high risk of bias for incomplete outcome data. Selective outcome reporting was assessed by evaluating differences in primary outcomes between publication and trial registration, if available. If no trial registration was available for a study, the study was coded as being at unknown risk for selective outcome reporting. We computed a summary score for each study by summing the number of items scoring high risk of bias, which could range from 0 to 5, with low scores indicating low risk of bias.

Meta-analysis

Meta-analyses were conducted to assess the effect of VRE-CBT on the severity of anxiety symptoms when compared with waitlist and CBT conditions. We calculated the effect sizes (Hedges *g*) for the difference between the conditions and their 95% CIs for posttest and follow-up measurements if data were available. For the follow-up date, we chose the longest follow-up period available. The effect sizes were based on intention-to-treat analysis, if available. Otherwise, completer analysis results were used. Heterogeneity among the studies was examined using the Cochran *Q* test and Higgins I^2 statistics [15]. The possibility of publication bias was evaluated by visual inspection of the funnel plot and Egger test [16] and using the Duval and Tweedie trim and fill procedure [17]. Random effects models were used as heterogeneity across the studies was expected, for example, because of variations in the number of sessions, treatment protocols, recruitment setting, and type of anxiety disorder.

Predefined subgroup analyses were conducted for type of anxiety disorder and method of recruitment (community, clinical, or both). Meta-regression analyses were used to estimate the impact of year of publication, quality of the individual studies (risk of bias), and treatment duration (in sessions). We conducted sensitivity analyses excluding outliers and studies in which the VRE-CBT condition included additional CBT elements or sessions with psychoeducation, cognitive restructuring, interoceptive exposure, and exposure

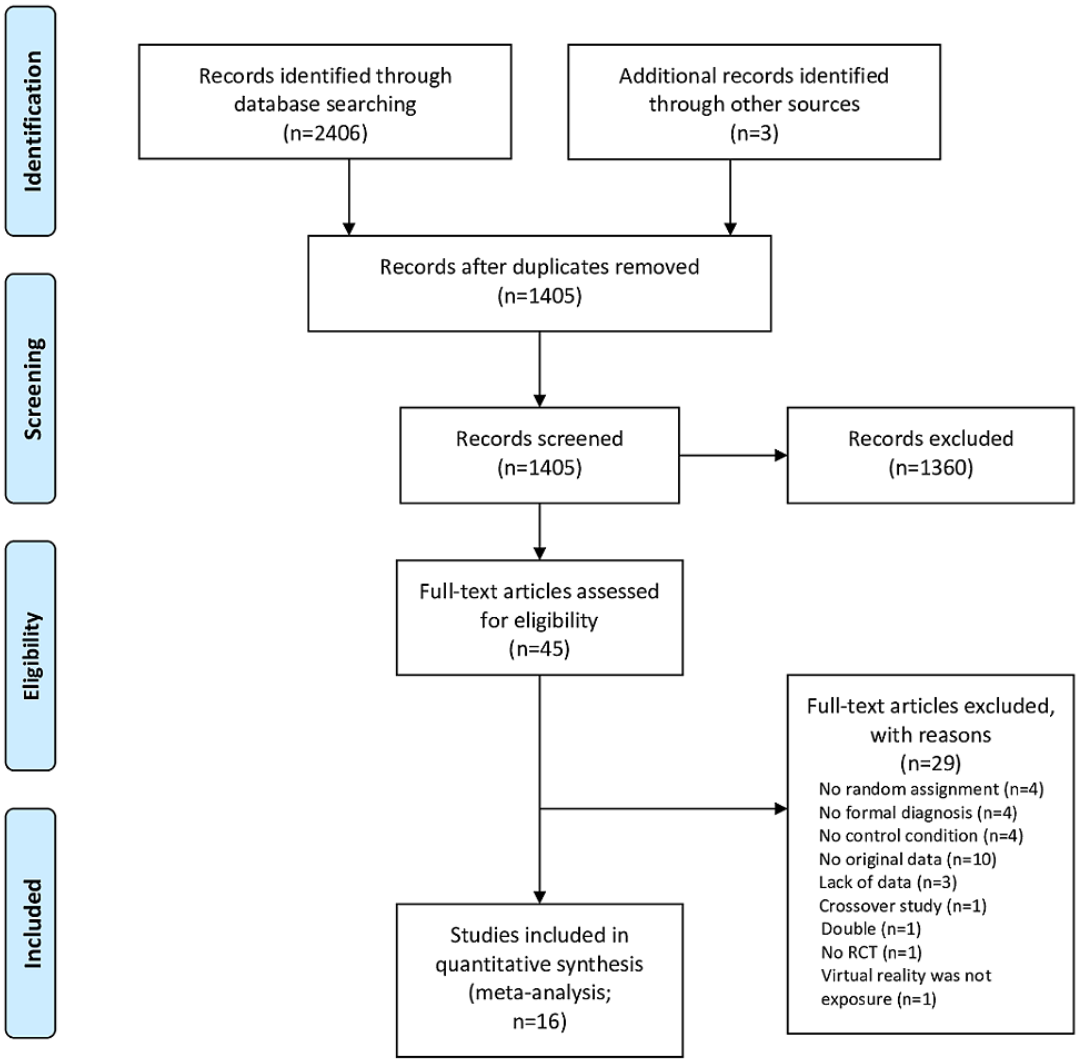
in vivo. A separate meta-analysis was performed to compare the dropouts between the VRE-CBT and CBT conditions to assess the feasibility of VRE-CBT. Odds ratios indicated the odds of participants dropping out from the VRE-CBT versus CBT. The analyses were performed using Comprehensive Meta-Analysis (version 2.2; Biostat Inc) [18].

Results

Study Inclusion

After removal of duplicates, the literature searches resulted in a total of 1405 papers for consideration. A total of 2 papers (2/1405, 0.1%) were identified from checking references. After screening the titles and abstracts, of the 1405 papers, 45 (3.2%) full-text articles were retrieved for a more detailed evaluation of eligibility. Figure 1 shows the flowchart of the inclusion process following the PRISMA guidelines [19]. Subsequently, of the 45 articles, 26 (58%) were excluded as they did not meet the inclusion criteria for various reasons. In addition, 3 studies (3/45, 7%) were excluded because of missing data for analyses despite efforts to acquire these data by contacting the authors (for specific reasons, see Figure 1). Finally, 16 studies were included in the meta-analysis.

Figure 1. Flowchart of study inclusion. RCT: randomized controlled trial.



Study Characteristics

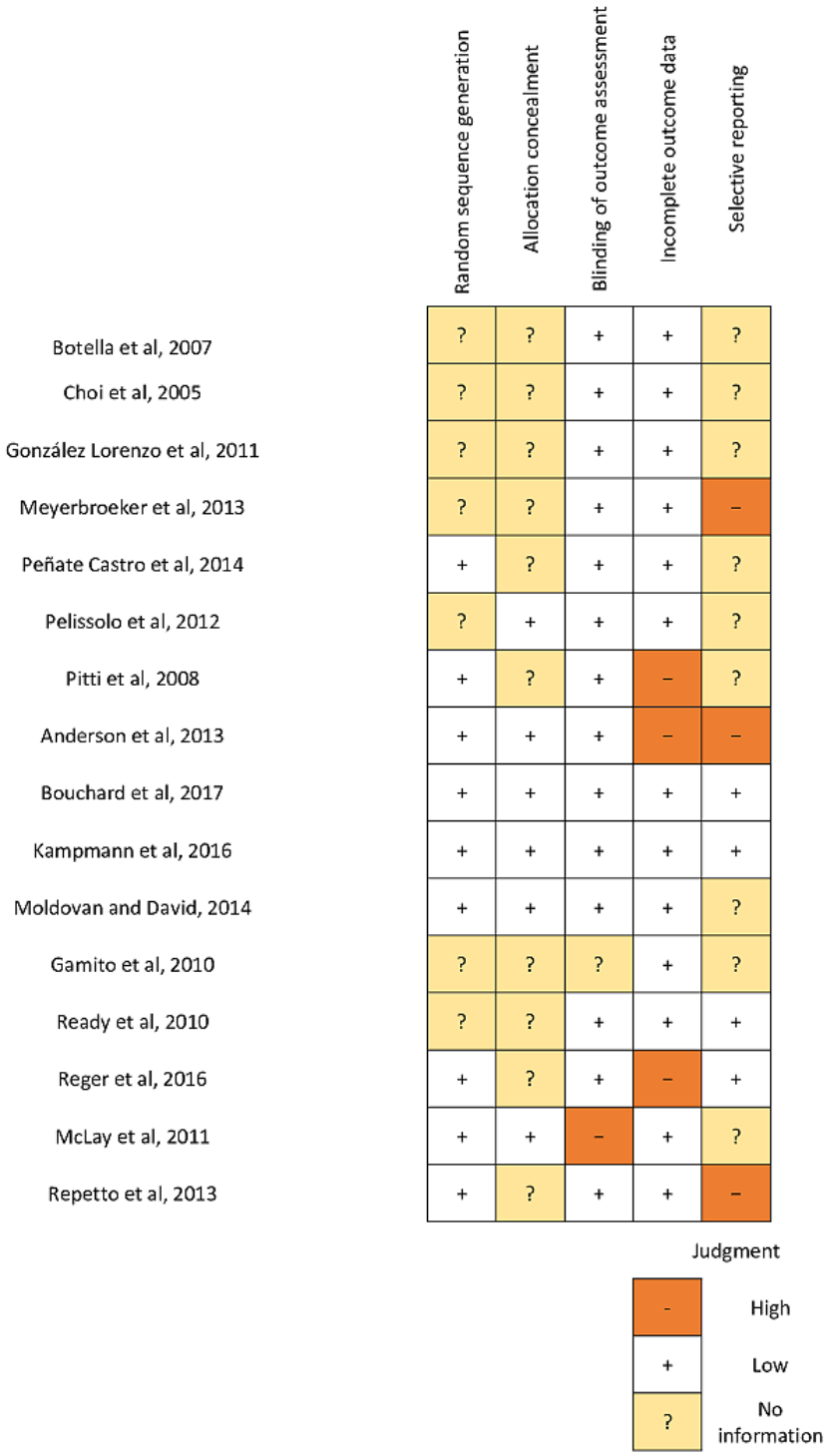
A total of 16 trials were included in Multimedia Appendix 2 [13,20-34]. These trials entailed 10 comparisons between VRE-CBT and a waitlist and 13 comparisons between VRE-CBT and a CBT condition. The number of participants per study ranged from 10 to 162 (12/16, 75% of the studies had <25 participants in each arm). Follow-up data were available for 38% (5/13) of the comparisons between VRE-CBT and a CBT condition and ranged from 6 to 12 months. Of the 16 studies, 7 (44%) focused on PD with or without AGO, 4 (25%) focused on SAD, 4 (25%) focused on PTSD, and 1 (6%) focused on GAD. None of the screened studies on OCD were included as they did not meet our inclusion criteria.

Risk of Bias

In each trial, 5 areas of risk of bias were rated as low, unclear (because of lack of information), or high risk of bias. As blinding of participants and personnel was not possible, all studies had a high risk of performance bias; therefore, this was not reported

(see Figure 2 for the risk of bias summary). Information on random sequence generation and allocation concealment was absent or not reported clearly in 44% (7/16) and 63% (10/16) of the trials, respectively. Of the 16 trials, 9 (56%) were scored with unclear risk of bias for selective outcome reporting as registration in a trial database could not be found.

Figure 2. Risk of bias summary [13,20-34].

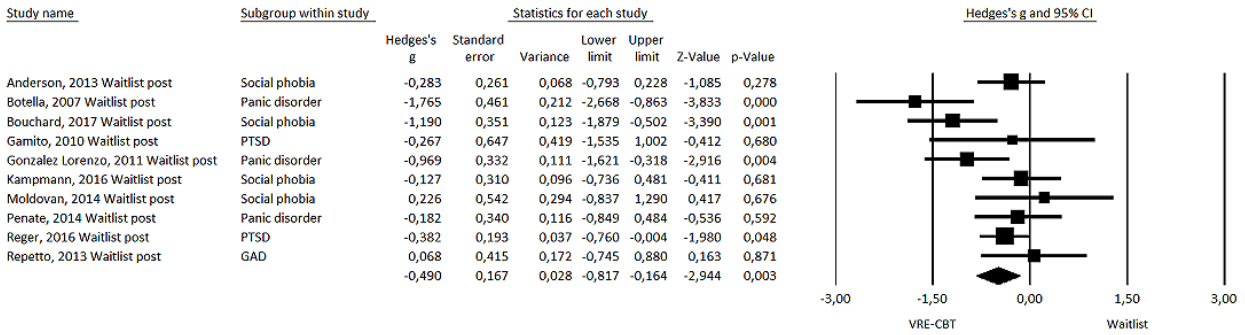


Effect Size: VRE-CBT Versus Waitlist at Posttest Measurement

The mean effect size of VRE-CBT on symptoms of anxiety when compared with waitlist control at posttest measurement (n_{co}=10) was medium and significant, favoring VRE-CBT (Hedges g=−0.490, 95% CI −0.82 to −0.16; P=.003; see Figure 3 for the forest plot). There was moderate heterogeneity (I^2=56%). Effect size decreased after exclusion of 1 potential outlier [20] (Hedges g=−0.391, 95% CI −0.66 to −0.12; P=.005),

with low heterogeneity (I^2=33%). Sensitivity analysis excluding studies in which VRE-CBT included additional CBT elements or sessions (n_{co}=6) showed a decrease in the effect size to Hedges *g* of −0.255 (95% CI −0.49 to −0.02; *P*=.03). Sensitivity analysis excluding 1 study with medication (antidepressants; n_{co}=9) as part of the intervention and control conditions showed a decrease in the effect size to Hedges *g* of −0.432 (95% CI −0.78 to −0.09; *P*=.01).

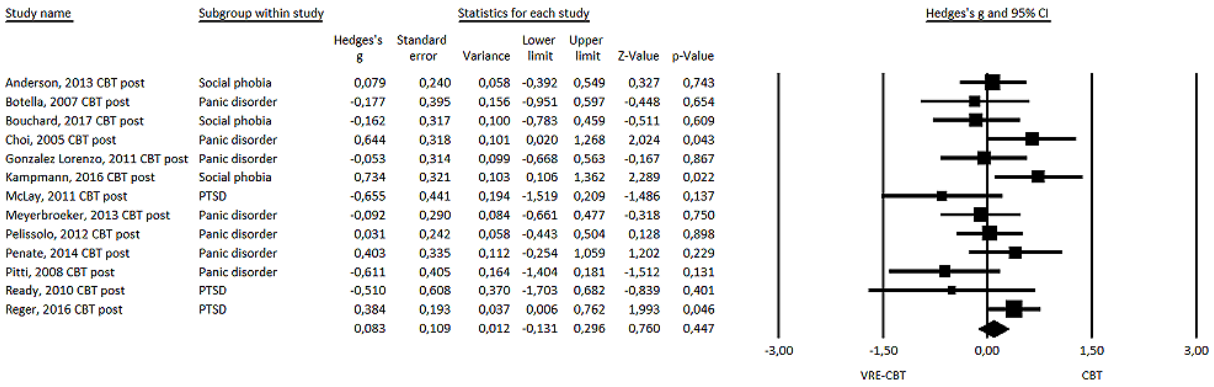
Figure 3. Forest plot of virtual reality exposure–based cognitive behavioral therapy versus waitlist at posttest measurement [20,22,24,26-30,32,34]. GAD: generalized anxiety disorder; PTSD: posttraumatic stress disorder; VRE-CBT: virtual reality exposure–based cognitive behavioral therapy.



Effect Size: VRE-CBT Versus CBT at Posttest Measurement

The mean effect size of VRE-CBT on symptoms of anxiety when compared with CBT at posttest measurement (n_{co}=13) was small and nonsignificant, favoring CBT (Hedges *g*=0.083, 95% CI −0.13 to 0.30; *P*=.45) with low to moderate heterogeneity (I^2=36%; see Figure 4 for the forest plot). Sensitivity analysis excluding studies in which VRE-CBT included additional CBT elements or sessions (n_{co}=6) showed a nonsignificant increase in the effect size to Hedges *g* of 0.209 (95% CI −0.14 to 0.56; *P*=.24). Sensitivity analysis excluding 1 study with medication (antidepressants; n_{co}=12) as part of the intervention and control conditions showed a similar effect size of Hedges *g* of 0.09 (95% CI −0.14 to 0.33; *P*=.34).

Figure 4. Forest plot of virtual reality exposure–based cognitive behavioral therapy versus cognitive behavioral therapy at posttest measurement [21-28,31-33]. CBT: cognitive behavioral therapy; PTSD: posttraumatic stress disorder; VRE-CBT: virtual reality exposure–based cognitive behavioral therapy.



Effect Size: VRE-CBT Versus CBT Condition at Follow-up (6-12 Months)

In total, there were 5 studies (5/16, 31%) with follow-up assessments that were compared with CBT. Of the 16 studies, 3 (19%) only conducted follow-up assessments after 6 months [24,27,31], 1 (6%) had 6- and 12-month assessments [25], and only 1 (6%) had a 12-month follow-up [20]. For the meta-analysis, the longest follow-up period was used. The mean effect size of VRE-CBT on symptoms of anxiety when compared with CBT (n_{co}=5) remained nonsignificant at follow-up (Hedges *g*=−0.082, 95% CI −0.40 to 0.24; *P*=.61; I^2=0).

Subgroup Analyses and Meta-Regression Analyses

Subgroup analyses were performed for the type of anxiety disorder and type of VRE-CBT (VRE either with or without regular CBT). Type of anxiety disorder (SAD: n_{co}=3; PD with or without AGO, or AGO: n_{co}=7; and PTSD: n_{co}=3) was a significant moderator (*P*=.005) for the comparison between VRE-CBT and waitlist, although this result was most likely affected by the high heterogeneity within 2 of the small

subgroups—PD with or without AGO, or AGO (I^2=75%) and SAD (I^2=59%). Only 1 GAD study (1/16, 6%) was available; therefore, we did not include this study in the analysis.

We were not able to conduct planned subgroup analyses for type of recruitment (community, clinical, or both) as there was only 1 study (1/16, 6%) with community recruitment, which did not allow for subgroup analysis.

Meta-regression to estimate the impact of year of publication (*P*=.16), treatment duration (in sessions; *P*=.85), and quality of the individual studies (risk of bias; *P*=.32) showed no significant effects in the comparison of VRE-CBT with waitlist control. In the comparison with CBT, year of publication (*P*=.44), treatment duration (in sessions; *P*=.07), and quality of the individual studies (risk of bias; *P*=.50) also showed no significant differences.

Comparison of Attrition Between VRE-CBT and CBT

A separate meta-analysis was conducted to compare the dropouts between VRE-CBT and CBT conditions to assess the feasibility of VRE-CBT. The dropout rates between the 2 conditions (n_{co}=10) showed no significant difference (odds ratio 0.79, 95% CI 0.49-1.27; *P*=.32).

Publication Bias

We conducted the Egger test for asymmetry of the funnel plot for the comparison between VRE-CBT and waitlist. Visual inspection of the funnel plot seemed to show some asymmetry, which could indicate small study effects. The Egger regression intercept test was not statistically significant (intercept=−0.53, 95% CI −4.09 to 3.03; *P*=.74). The Duval and Tweedie trim and fill procedure showed that adjustment for potentially missing studies (n=1) was associated with the effect size slightly increasing from −0.51 (95% CI −0.84 to −0.17) to −0.56 (95% CI −0.88 to −0.22; Figure 5). The Egger test for the comparison between VRE with and without CBT also showed some asymmetry. The Egger regression intercept test was not statistically significant (intercept=−2.15, 95% CI −4.60 to 0.30; *P*=.08). The Duval and Tweedie trim and fill procedure showed that adjustment for potentially missing studies (n=3) was associated with the effect size increasing from 0.09 (95% CI −0.13 to 0.30) to 0.19 (95% CI −0.04 to 0.42; Figure 6), with largely overlapping CIs.

Figure 5. Trim and fill adjusted funnel plot of virtual reality exposure–based cognitive behavioral therapy versus waitlist (the white circles represent the observed studies, and the black circles represent the imputed studies). Std diff: Standard difference.

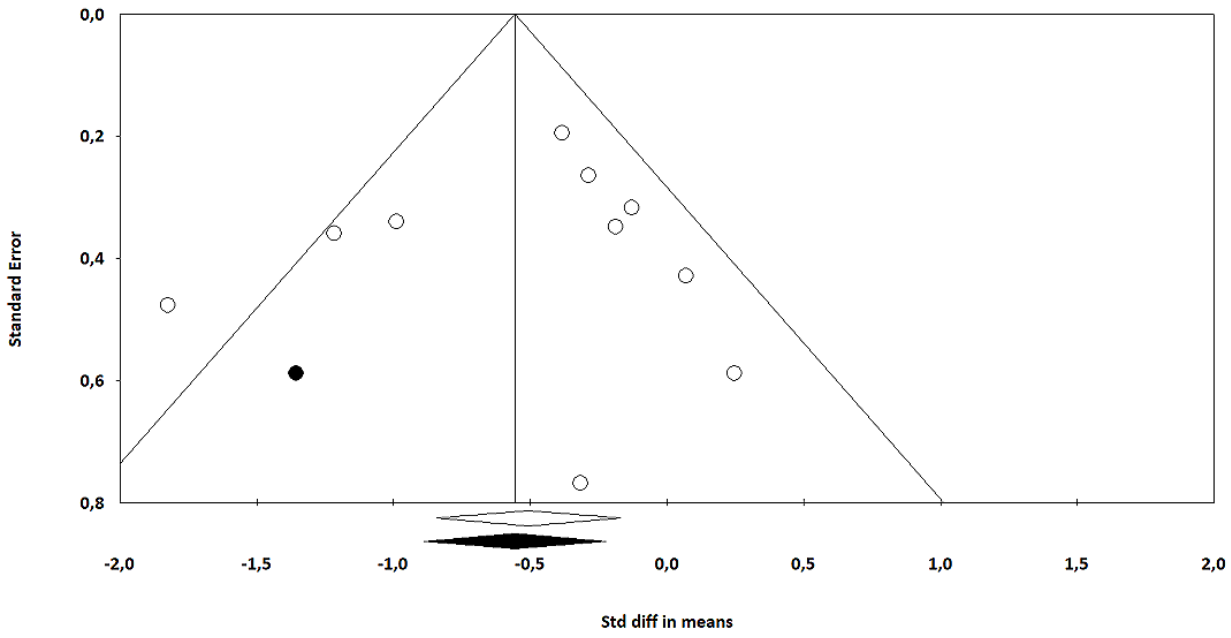
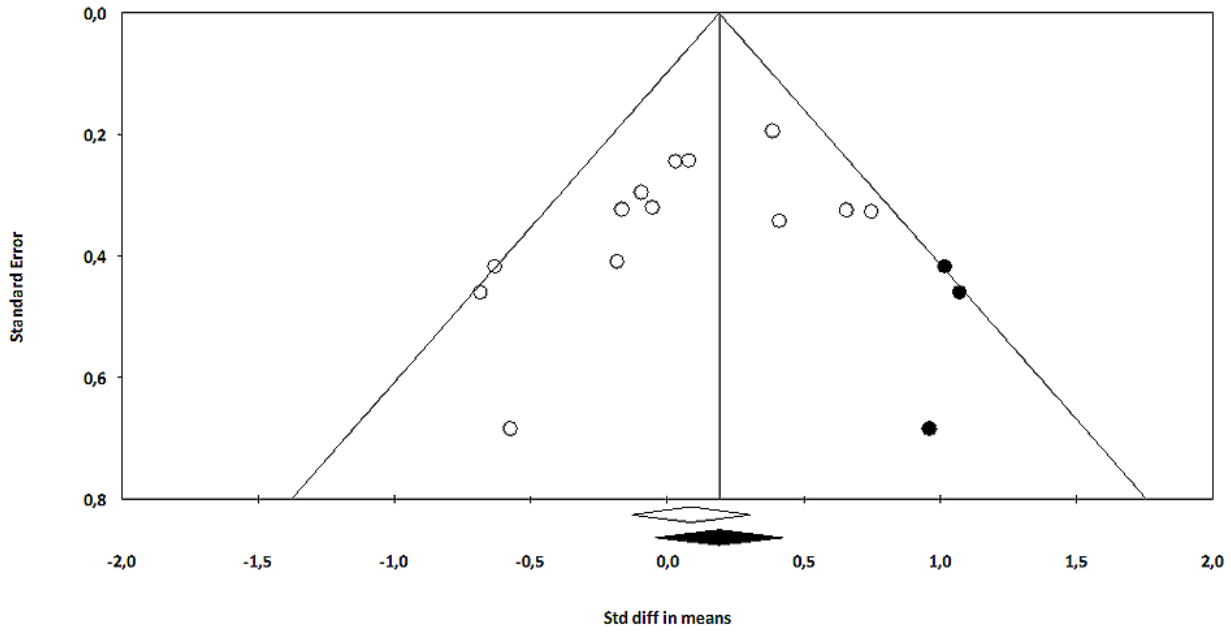


Figure 6. Trim and fill adjusted funnel plot of virtual reality exposure–based cognitive behavioral therapy versus cognitive behavioral therapy (the white circles represent the observed studies, and the black circles represent the imputed studies). Std diff: Standard difference.



Discussion

Principal Findings

This is the first meta-analysis of VRE-CBT for anxiety disorders focusing solely on patients with more severe anxiety disorders with a formal diagnosis, excluding specific phobias. The meta-analysis, consisting of 16 VRE-CBT trials including 817 patients with anxiety disorders, showed a medium effect size compared with waitlist conditions. In the comparison between VRE-CBT and CBT, no significant differences were found between posttest and follow-up measurements. These findings are somewhat less positive than findings from previous meta-analyses as the effect sizes of VRE-CBT were lower in our meta-analysis. We found an effect size (Hedges *g*) of −0.49 of VRE-CBT versus waitlist compared with an estimated Hedges *g* of 0.88 and 0.79 in the meta-analyses by Carl et al [9] and Fodor et al [10], respectively. It can be hypothesized that their higher effect size is a result of the inclusion of specific phobias and subthreshold anxiety disorders. Specific phobias seemed to account for the larger effect sizes in the meta-analysis by Fodor et al [10], in which high effects of VRE-CBT for specific phobias and fight anxiety, and moderate or small effects for SAD and PTSD were found. However, in the study by Carl et al [9], this was not the case. Thus, the influence of specific phobias on the effect size is inconclusive. With regard to subthreshold anxiety disorders, neither study requested that participants be formally diagnosed with anxiety disorders. Therefore, it can be assumed that less severe cases were also included. It could be argued that people with less severe anxiety are easier to treat, and a greater effect can be reached in treatment. On the other hand, the reverse may also be true—with more symptoms at the start of treatment, greater gains can be made. Fodor et al [10] conducted sensitivity analyses that were restricted to patients with an anxiety disorder diagnosed with a clinical interview or a cutoff score on a questionnaire, which resulted in a smaller effect (Hedges *g*=0.72). This finding seems to support the hypothesis that people with less severe disorders might benefit more from treatment with VRE-CBT. With regard to the role of CBT elements or CBT sessions in 56% (9/16) of the VRE-CBT studies, sensitivity analyses with only pure VRE studies showed that pure VRE studies without additional CBT elements showed lower effect sizes in the comparison with the waitlist condition, indicating that adding CBT elements (such as psychoeducation, cognitive restructuring, interoceptive exposure, and exposure in vivo) to VRE is advised for optimizing effects.

Compared with the other studies, it is more difficult to interpret the effect size of the only study on GAD [34] in this meta-analysis as the VRE-CBT in this study mostly consisted of progressive muscle relaxation and autogenic techniques, and only 2 of 8 sessions consisted of exposure therapy. Furthermore, exposure itself differed from regular exposure-based interventions in GAD. It involved exposure to preselected words or images related to personal stressful events, although in regular CBT treatments patients with GAD are exposed to worries about hypothetical scenarios describing their worst fear [35]. It is possible that more focused exposure could lead to better results in future VRE-CBT studies for GAD.

We found no difference in attrition between VRE-CBT and CBT, which is in line with the findings of Benbow and Anderson [4], who conducted a meta-analysis on attrition between VRE-CBT and CBT across 46 studies with a combined sample size of 1057 participants. Fodor et al [10] also found that VRE-CBT yielded similar dropout rates to other interventions. These findings also suggest that VRE-CBT does not mitigate attrition.

The heterogeneity of the studies was low to moderate and did not seem to influence the results. We found no indications of small study effects or publication bias.

The risk of bias assessment in our study showed that, in many studies, it was difficult to assess risk of bias because of lacking or unclear information on random sequence generation and allocation concealment and because of lack of registration in a trial register. Keeping this in mind, it is difficult to interpret our results from the meta-regression, in which we found no significant effect of the total risk of bias score. Fodor et al [10] also used the Cochrane tool [14] for assessing the risk of bias in their meta-analysis on VRE-CBT in anxiety and depression, and they found that most studies had high or uncertain risk of bias across domains. In exploratory subgroup analyses, they also did not find differences between studies with high, uncertain, and low risk of bias, but these results must be interpreted with caution, as there were only few studies with a small risk of bias. Furthermore, previous research on adherence to study quality criteria in VRE-CBT studies on anxiety disorders by McCann et al [36] showed that the studies met an average of 2.85 (SD 1.56) of 8 quality criteria for research design, but study quality did not affect the effect size. It may be concluded that better reporting of quality in VRE-CBT RCTs is needed before conclusions can be drawn on the effect of quality on effect sizes.

Limitations

This meta-analysis has some limitations. First, there were a limited number of trials per diagnosis (ie, none on OCD and only 1/16, 6% on GAD) and per method of recruitment (only 1/16, 6% recruited in the community) and few studies with follow-up outcome data (5/16, 31%). Subgroup analyses showed that *type of anxiety disorder* was a significant moderator in the waitlist comparison, but no conclusions can be drawn from this result, as there was high heterogeneity within the subgroups, expressing excessive clinical diversity and therefore showing no true effect. Second, we did not investigate the sense of presence or other indicators of interaction, although presence may influence the effectiveness of VRE-CBT. Ling et al [37] found a positive relationship between sense of presence and anxiety disorders, which was stronger in studies with participants with a formal anxiety disorder than in nonclinical populations. Third, 3 studies (3/45, 7%) were excluded because of missing data for analyses despite efforts to acquire these data by contacting the authors. This may have influenced the estimations.

Clinical Implications

There are some clinical implications of the findings in this meta-analysis and for the position that VRE-CBT holds in the treatment of more severe anxiety disorders. Patients should be made aware of the state of the current evidence regarding VRE-CBT relative to alternatives to be able to make a shared decision with the therapist. The advantages of VRE-CBT over CBT may play a role in this decision. Bouchard et al [27] used a Specific Work for Exposure Applied in Therapy scale (evaluating topics such as cost, time, planning, and difficulties), which showed that VRE-CBT was more cost-efficient and practical for therapists than CBT with in vivo exposure. In addition, VRE-CBT offers a large potential for exposing patients to situations and stimuli that are too complex, costly, or challenging for in vivo exercises. It can be hypothesized that when these applications are exploited accordingly, it creates the potential to increase the effectiveness of VRE-CBT above and beyond in vivo exposure.

Furthermore, there are some specific concerns regarding VRE-CBT that clinicians should take into account: the use of the VR simulators can cause simulator sickness, appropriate training and supervision are needed, and technical support should be available in case of technical problems. In addition, although lower-cost options are becoming available, such as using smartphones as VR viewers, they still need to be evaluated [38]. Regarding improving VRE-CBT, developing VR that can be customized precisely to a patient's needs with a better sense of presence and interaction may lead to even better outcomes.

Conclusions

In conclusion, the results of our study show that VRE-CBT is as effective as CBT in the treatment of more severe formal anxiety disorders. However, VRE-CBT studies in this meta-analysis not only were limited in number but the quality of reporting was also poor. Therefore, awaiting further high-quality data, VRE-CBT may be considered a promising alternative to CBT, especially for patients who prefer VRE-CBT over CBT [5].

Future research should focus on conducting high-quality RCTs. Subsequently, examining for which patients or anxiety disorders VRE-CBT or CBT works best could be a next step in research.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strings for PubMed, PsycINFO, and Embase.
[DOCX File , 15 KB - jmir_v24i2e26736_app1.docx]

Multimedia Appendix 2
Study characteristics.
[DOCX File , 30 KB - jmir_v24i2e26736_app2.docx]

References

1. Olatunji BO, Cisler JM, Deacon BJ. Efficacy of cognitive behavioral therapy for anxiety disorders: a review of meta-analytic findings. Psychiatr Clin North Am 2010;33(3):557-577. [doi: 10.1016/j.psc.2010.04.002] [Medline: 20599133]
2. Carpenter JK, Andrews LA, Witcraft SM, Powers MB, Smits JAJ, Hofmann SG. Cognitive behavioral therapy for anxiety and related disorders: a meta-analysis of randomized placebo-controlled trials. Depress Anxiety 2018;35(6):502-514 [FREE Full text] [doi: 10.1002/da.22728] [Medline: 29451967]
3. Maples-Keller JL, Bunnell BE, Kim S, Rothbaum BO. The use of virtual reality technology in the treatment of anxiety and other psychiatric disorders. Harv Rev Psychiatry 2017;25(3):103-113. [doi: 10.1097/hrp.0000000000000138] [Medline: 28475502]
4. Benbow AA, Anderson PL. A meta-analytic examination of attrition in virtual reality exposure therapy for anxiety disorders. J Anxiety Disord 2019;61:18-26. [doi: 10.1016/j.janxdis.2018.06.006] [Medline: 30646997]
5. Garcia-Palacios A, Botella C, Hoffman H, Fabregat S. Comparing acceptance and refusal rates of virtual reality exposure vs. in vivo exposure by patients with specific phobias. Cyberpsychol Behav 2007;10(5):722-724. [doi: 10.1089/cpb.2007.9962] [Medline: 17927544]
6. Opriş D, Pintea S, García-Palacios A, Botella C, Szamosközi Ş, David D. Virtual reality exposure therapy in anxiety disorders: a quantitative meta-analysis. Depress Anxiety 2012;29(2):85-93. [doi: 10.1002/da.20910] [Medline: 22065564]
7. Morina N, Ijntema H, Meyerbröker K, Emmelkamp PM. Can virtual reality exposure therapy gains be generalized to real-life? A meta-analysis of studies applying behavioral assessments. Behav Res Ther 2015;74:18-24. [doi: 10.1016/j.brat.2015.08.010] [Medline: 26355646]
8. Powers M, Emmelkamp PM. Virtual reality exposure therapy for anxiety disorders: a meta-analysis. J Anxiety Disord 2008;22(3):561-569. [doi: 10.1016/j.janxdis.2007.04.006] [Medline: 17544252]
9. Carl E, Stein A, Levihn-Coon A, Pogue JR, Rothbaum B, Emmelkamp P, et al. Virtual reality exposure therapy for anxiety and related disorders: a meta-analysis of randomized controlled trials. J Anxiety Disord 2019;61:27-36. [doi: 10.1016/j.janxdis.2018.08.003] [Medline: 30287083]
10. Fodor LA, Cote CD, Cuijpers P, Szamoskozi Ş, David D, Cristea IA. The effectiveness of virtual reality based interventions for symptoms of anxiety and depression: a meta-analysis. Sci Rep 2018;8(1):10323 [FREE Full text] [doi: 10.1038/s41598-018-28113-6] [Medline: 29985400]
11. Botella C, Fernández-Álvarez J, Guillén V, García-Palacios A, Baños R. Recent progress in virtual reality exposure therapy for phobias: a systematic review. Curr Psychiatry Rep 2017;19(7):42. [doi: 10.1007/s11920-017-0788-4] [Medline: 28540594]
12. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
13. Pitti C, Peñate W, de la Fuente J, Bethencourt J, Acosta L, Villaverde M, et al. [Agoraphobia: combined treatment and virtual reality. Preliminary results]. Actas Esp Psiquiatr 2008;36(2):94-101 [FREE Full text] [Medline: 18365789]
14. Higgins JP, Altman DG. Assessing risk of bias in included studies. In: Higgins JP, Green S, editors. Cochrane handbook for systematic reviews of interventions. Hoboken: John Wiley & Sons; 2008:187-241.
15. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327(7414):557-560 [FREE Full text] [doi: 10.1136/bmj.327.7414.557] [Medline: 12958120]
16. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315(7109):629-634. [doi: 10.1136/bmj.315.7109.629] [Medline: 9310563]
17. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000;56(2):455-463. [doi: 10.1111/j.0006-341x.2000.00455.x] [Medline: 10877304]
18. Borenstein M, Hedges L, Higgins JP, Rothstein HR. Comprehensive meta-analysis (Version 2.2.027) [Computer software]. Comprehensive meta-analysis. Englewood: Biostat; 2005. URL: https://www.meta-analysis.com/ [accessed 2022-01-20]
19. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 2009;6(7):e1000100. [doi: 10.1371/journal.pmed.1000100] [Medline: 19621070]
20. Botella C, García-Palacios A, Villa H, Baños RM, Quero S, Alcañiz M, et al. Virtual reality exposure in the treatment of panic disorder and agoraphobia: a controlled study. Clin Psychol Psychother 2007;14(3):164-175. [doi: 10.1002/cpp.524]
21. Choi Y, Vincelli F, Riva G, Wiederhold BK, Lee JH, Park KH. Effects of group experiential cognitive therapy for the treatment of panic disorder with agoraphobia. Cyberpsychol Behav 2005;8(4):387-393. [doi: 10.1089/cpb.2005.8.387] [Medline: 16092896]
22. Gonzalez LM, Peñate W, Pitti C, Bethencourt-Pére JM, de la Fuente Portero J, Marco RG. Efficacy of virtual reality exposure therapy combined with two pharmacotherapies in the treatment of agoraphobia. Int J Clin Heal Psychol 2011;11(2):189-203.
23. Meyerbroeker K, Morina N, Kerkhof GA, Emmelkamp PM. Virtual reality exposure therapy does not provide any additional value in agoraphobic patients: a randomized controlled trial. Psychother Psychosom 2013;82(3):170-176. [doi: 10.1159/000342715] [Medline: 23548832]

24. Peñate Castro W, Roca Sánchez MJ, Pitti González CT, Bethencourt JM, de la Fuente Portero JA, Marco RG. Cognitive-behavioral treatment and antidepressants combined with virtual reality exposure for patients with chronic agoraphobia. Int J Clin Health Psychol 2014;14(1):9-17. [doi: 10.1016/s1697-2600(14)70032-8]
25. Pelissolo A, Zaoui M, Aguayo G, Yao SN, Roche S, Ecochard R, et al. Virtual reality exposure therapy versus cognitive behavior therapy for panic disorder with agoraphobia: a randomized comparison study. J Cybertherapy Rehabil 2012;5(1):35-43.
26. Anderson PL, Price M, Edwards SM, Obasaju MA, Schmertz SK, Zimand E, et al. Virtual reality exposure therapy for social anxiety disorder: a randomized controlled trial. J Consult Clin Psychol 2013;81(5):751-760. [doi: 10.1037/a0033559] [Medline: 23796315]
27. Bouchard S, Dumoulin S, Robillard G, Guitard T, Klinger É, Forget H, et al. Virtual reality compared with in vivo exposure in the treatment of social anxiety disorder: a three-arm randomised controlled trial. Br J Psychiatry 2017;210(4):276-283. [doi: 10.1192/bjp.bp.116.184234] [Medline: 27979818]
28. Kampmann IL, Emmelkamp PM, Hartanto D, Brinkman WP, Zijlstra BJ, Morina N. Exposure to virtual social interactions in the treatment of social anxiety disorder: a randomized controlled trial. Behav Res Ther 2016;77:147-156. [doi: 10.1016/j.brat.2015.12.016] [Medline: 26752328]
29. Moldovan R, David D. One session treatment of cognitive and behavioral therapy and virtual reality for social and specific phobias. Preliminary results from a randomized clinical trial. J Evid Based Psychother 2014;14(1):67-83.
30. Gamito P, Oliveira J, Rosa P, Morais D, Duarte N, Oliveira S, et al. PTSD elderly war veterans: a clinical controlled pilot study. Cyberpsychol Behav Soc Netw 2010;13(1):43-48. [doi: 10.1089/cyber.2009.0237] [Medline: 20528292]
31. Ready DJ, Gerardi RJ, Backscheider AG, Mascaro N, Rothbaum BO. Comparing virtual reality exposure therapy to present-centered therapy with 11 U.S. Vietnam veterans with PTSD. Cyberpsychol Behav Soc Netw 2010;13(1):49-54. [doi: 10.1089/cyber.2009.0239] [Medline: 20528293]
32. Reger GM, Koenen-Woods P, Zetocha K, Smolenski DJ, Holloway KM, Rothbaum BO, et al. Randomized controlled trial of prolonged exposure using imaginal exposure vs. virtual reality exposure in active duty soldiers with deployment-related posttraumatic stress disorder (PTSD). J Consult Clin Psychol 2016;84(11):946-959. [doi: 10.1037/ccp0000134] [Medline: 27606699]
33. McLay RN, Wood DP, Webb-Murphy JA, Spira JL, Wiederhold MD, Pyne JM, et al. A randomized, controlled trial of virtual reality-graded exposure therapy for post-traumatic stress disorder in active duty service members with combat-related post-traumatic stress disorder. Cyberpsychol Behav Soc Netw 2011;14(4):223-229. [doi: 10.1089/cyber.2011.0003] [Medline: 21332375]
34. Repetto C, Gaggioli A, Pallavicini F, Cipresso P, Raspelli S, Riva G. Virtual reality and mobile phones in the treatment of generalized anxiety disorders: a phase-2 clinical trial. Pers Ubiquit Comput 2013;17(2):253-260. [doi: 10.1007/s00779-011-0467-0]
35. Dugas MJ, Brillon P, Savard P, Turcotte J, Gaudet A, Ladouceur R, et al. A randomized clinical trial of cognitive-behavioral therapy and applied relaxation for adults with generalized anxiety disorder. Behav Ther 2010;41(1):46-58 [FREE Full text] [doi: 10.1016/j.beth.2008.12.004] [Medline: 20171327]
36. McCann RA, Armstrong CM, Skopp NA, Edwards-Stewart A, Smolenski DJ, June JD, et al. Virtual reality exposure therapy for the treatment of anxiety disorders: an evaluation of research quality. J Anxiety Disord 2014;28(6):625-631. [doi: 10.1016/j.janxdis.2014.05.010] [Medline: 25093964]
37. Ling Y, Nefs HT, Morina N, Heynderickx I, Brinkman WP. A meta-analysis on the relationship between self-reported presence and anxiety in virtual reality exposure therapy for anxiety disorders. PLoS One 2014;9(5):e96144 [FREE Full text] [doi: 10.1371/journal.pone.0096144] [Medline: 24801324]
38. Mishkind MC, Norr AM, Katz AC, Reger GM. Review of virtual reality treatment in psychiatry: evidence versus current diffusion and use. Curr Psychiatry Rep 2017;19(11):80. [doi: 10.1007/s11920-017-0836-0] [Medline: 28920179]

Abbreviations

AGO: agoraphobia
CBT: cognitive behavioral therapy
DSM: Diagnostic and Statistical Manual of Mental Disorders
GAD: generalized anxiety disorder
ICD: International Classification of Diseases
OCD: obsessive-compulsive disorder
PD: panic disorder
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PTSD: posttraumatic stress disorder
SAD: social anxiety disorder
VR: virtual reality
VRE: virtual reality exposure

VRE-CBT: virtual reality exposure–based cognitive behavioral therapy

Edited by A Mavragani; submitted 23.12.20; peer-reviewed by S Bouchard, C Jonassaint, R Mandryk; comments to author 31.03.21; revised version received 22.05.21; accepted 06.12.21; published 10.02.22.

Please cite as:
van Loenen I, Scholten W, Muntingh A, Smit J, Batelaan N
The Effectiveness of Virtual Reality Exposure–Based Cognitive Behavioral Therapy for Severe Anxiety Disorders, Obsessive-Compulsive Disorder, and Posttraumatic Stress Disorder: Meta-analysis
J Med Internet Res 2022;24(2):e26736
URL: https://www.jmir.org/2022/2/e26736
doi:10.2196/26736
PMID:35142632

Review

The Effect of Online Health Information Seeking on Physician-Patient Relationships: Systematic Review

Aijing Luo[1,2,3], PhD; Lu Qin[2,3], MS; Yifeng Yuan[1,2], MS; Zhengzijin Yang[2,3], MS; Fei Liu[2,3], MS; Panhao Huang[4], PhD; Wenzhao Xie[2,3], PhD

[1]The Second Xiangya Hospital, Central South University, Changsha, China

[2]Key Laboratory of Medical Information Research, The Third Xiangya Hospital, Central South University, Changsha, China

[3]School of Life Sciences, Central South University, Changsha, China

[4]Department of Pharmacy, The Third Xiangya Hospital, Central South University, Changsha, China

Corresponding Author:
Wenzhao Xie, PhD
Key Laboratory of Medical Information Research
The Third Xiangya Hospital
Central South University
No.138 Tongzipo Road
Changsha, 410013
China
Phone: 86 0731 8861 8316
Email: xie_wenzhao@126.com

Abstract

Background: The internet has now become part of human life and is constantly changing people's way of life. With the increasing popularity of online health information (OHI), it has been found that OHI can affect the physician-patient relationship by influencing patient behaviors.

Objective: This study aims to systematically investigate the impact of OHI-seeking behavior on the physician-patient relationship.

Methods: Literature retrieval was conducted on 4 databases (Web of Science, PubMed, China National Knowledge Infrastructure, SinoMed), and the time limit for literature publication was before August 1, 2021.

Results: We selected 53 target papers (42 [79%] English papers and 11 [21%] Chinese papers) that met the inclusion criteria. Of these, 31 (58%) papers believe that patients' OHI behavior can enable them to participate in their own medical care, improve patient compliance, and improve the physician-patient relationship. In addition, 14 (26%) papers maintain a neutral attitude, some believing that OHI behavior has no significant effect on doctors and patients and others believing that due to changes in the factors affecting OHI behavior, they will have a negative or a positive impact. Furthermore, 8 (15%) papers believe that OHI search behavior has a negative impact on doctors and patients, while 6 (11%) papers show that OHI reduces Chinese patients' trust in doctors.

Conclusions: Our main findings showed that (1) OHI-seeking behavior has an impact on patients' psychology, behavior, and evaluation of doctors; (2) whether patients choose to discuss OHI with doctors has different effects on the physician-patient relationship; and (3) the negative impact of OHI on China's internet users is worthy of attention. Due to the low quality of OHI, poor health information literacy, short physician-patient communication time, and various types of negative news, patients' trust in doctors has declined, thus affecting the physician-patient relationship. Improvement of people's health information literacy and the quality of OHI are important factors that promote the positive impact of OHI on the physician-patient relationship.

(J Med Internet Res 2022;24(2):e23354) doi:10.2196/23354

KEYWORDS

online health information; search behavior; physician-patient relationship; physician-patient consultation.

Introduction

About 4.66 billion people worldwide have access to the internet [1]. The internet has gradually become part of human life, constantly changing people's lifestyle. As the availability and immediacy of information services provided by the internet continue to improve, and patients' private information can be concealed to a certain extent, internet health information services have become increasingly attractive [2]. In addition, the lack of medical resources makes people choose to obtain health information online to meet their own health information needs [3]. The online health information (OHI) that patients search for on the internet mainly includes information about diseases, nutrition, treatments, physical and mental health, etc [4,5]. The uneven quality of OHI has a major impact on patients. The credibility of information perceived by patients affects whether the patients use the internet as a frequently used and preferred information source [6]. Physicians are still the most popular source of health information, but the internet has gradually become another important source of health information [7].

OHI seeking can influence physician-patient relationships and patient compliance. Patients who can obtain more health information can better follow the treatment process and enjoy better therapeutic effects [8]. The rapid development of the internet has changed the access of patients to health information and affected the existing physician-patient relationship, which to a large extent determines the medical result that patients receive [9]. As mentioned before, patients choose to obtain health information online to meet their own needs. At the same time, the health information search results provided by the internet show that extreme situations, for example, advice that is contrary to the standard medical opinion or complex data provided by health care professionals, leads to misinterpretation, confusion, and other problems for patients [10]. Of course, the availability of health information on the internet is transforming many patients from passive medical service consumers to those who can participate in the medical process, which brings new challenges for many physicians [11,12]. When patients carry a lot of health information in a short consultation, can physicians deal with it as usual?

With the continuous development of the internet, the physician-patient relationship has attracted much attention. In terms of importance, the relationship between patients and physicians is second only to that of family [13]. It is viewed as extremely or very important by 67%, exceeding relationships with spiritual advisors, pharmacists, coworkers, and financial advisors [13]. Due to the patients' lack of understanding of diseases and communication barriers between physicians and patients, some patients cannot understand the results of diagnosis and treatment of the disease and the treatment behavior, which causes a series of problems [14]. Physician-patient communication is a complex clinical behavior whose main goal is to share medical information to improve the education of clinical diagnosis, treatment, and specific diseases [14]. The quality of physician-patient communication affects the physician-patient relationship. In the past, physicians made decisions and patients obeyed them, which constitutes the traditional physician-patient relationship [15]. Patients and medical staff advocate the transition to mutual participation, that is, shared power and responsibility [15]. Previous studies have shown that processing patients' OHI-seeking behavior in daily consultation can improve the quality of medical services [16]. In an ideal physician-patient relationship, patients should be guided instead of looking for OHI independently [17]. However, at present, patients are mainly looking for OHI by themselves, and they are unable to control the quality of information and other aspects.

In China, the total population is about 1.4 billion [18]. As of June 2021, the number of internet users reached 1.011 billion, and the internet penetration rate reached 71.6%. The “Healthy China” strategy is China’s priority development strategy [19]. By implementing internet medicine, the Healthy China strategy promotes the mobility of medical services, enhances the operation efficiency of the overall medical and health system, and optimizes the allocation of medical resources [19]. Due to China's large population and the impact of COVID-19, China's demand for medical resources is growing exponentially. Increasingly more doctors and patients are seeking health information through internet platforms, effectively breaking the time and space restrictions and giving China's unbalanced medical resources a chance to be redistributed [20]. Considering that OHI may have a positive or a negative impact on the physician-patient relationship, which is important in medical care, this study aims to examine the impact of OHI on the physician-patient relationship in China.

In recent years, studies on health information seeking have been increasing. It is of great significance to understand the impact of the current health information seeking on the physician-patient relationship. Thus, the purpose of this study is to systematically review the current studies on the impact of OHI seeking on the physician-patient relationship.

Methods

Literature Retrieval

In this study, English references were obtained from the databases Web of Science and PubMed. The PubMed database contains references related to medicine and life sciences. Web of Science includes the most influential core academic journals on natural science, engineering technology, biomedicine, and other research fields. Chinese references were obtained from the databases China National Knowledge Infrastructure (CNKI) and SinoMed. The CNKI is China's largest full-text journal database, while SinoMed focuses on collecting the biomedical literature in China.

After consulting with librarians, the search strategy for this paper consisted of all possible keywords related to 4 topics: (1) online OR internet OR web OR network, (2) wellness information OR health information, (3) search* OR seek* OR inquiry OR query, and (4) physician-patient communication OR doctor-patient communication OR physician-patient relation* OR doctor-patient relation* OR physician-patient interaction OR doctor-patient interaction OR physician-patient trust OR doctor-patient trust. The papers were published before August 31, 2021. Combinations of these keywords were

searched in the 4 databases to make the literature search as comprehensive as possible. In addition, we searched the PubMed database separately with Medical Subject Headings (MESH). The Mesh terms were “patient-physician relations” and “internet.”

This systematic review conforms to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Multimedia Appendix 1).

Inclusion and Exclusion Criteria of Publication Standards

To make the coverage of this study comprehensive enough, the types of papers included were journal papers, conference papers, and academic dissertations. Paper retrieval covered all regions and languages, but only papers with full text in English or Chinese were retained. Papers involving only OHI-seeking studies or only physician-patient relationship studies were excluded. We also excluded all nonempirical research papers, including reviews, research on websites, and research commentaries. Then, we evaluated the quality of the included papers. We used the Critical Appraisal Skills Programme quality assessment tool for qualitative studies, which comprises 10 questions [21] (Multimedia Appendix 2). We also used a quality assessment tool for quantitative studies that comprises 14 questions customized by Tan et al [22] (Multimedia Appendix 3). Papers with a quality assessment score lower than 0.7 were excluded.

Paper-Screening Process

The literature screening in this study was independently carried out by 2 researchers. They screened the titles and abstracts, respectively, and read the full text to extract opinions. We compared the 2 researchers' screening results and the consistency of the extracted views, discussed the discrepancies to ensure the consistency and integrity of results, and used quality assessment tools to assess the quality of the papers. Endnote 20 was used to merge related search results and delete duplicate papers.

Data Extraction and Management

The research data of papers were independently extracted by 2 researchers according to the predesigned table. It mainly included the following information: country, research design method, sample size, respondents, and conclusion. Where there was ambiguity, the 2 researchers discussed it and reached an agreement.

Results

Characteristics of the Papers

In this study, we searched the PubMed and Web of Science databases and retrieved 10,303 and 9345 records, respectively, for a total of 19,648 initially searched records and 15,801 (80.42%) exported records that remained after duplication removal using Endnote 20. According to the screening criteria, whether the papers discussed OHI seeking and the physician-patient relationship, 173 (1.09%) papers were included according to the title and abstract for further screening. Through the application of inclusion and exclusion criteria, of these 173 papers, 72 (41.6%) did not involve the impact of the physician-patient relationship; 7 (4%) were reviews; for 9 (5.2%), the original text could not be obtained; 19 (10.9%) did not have health information seeking as the main research object; 13 (7.5%) had full text in languages other than English; 10 (5.8%) were on nonempirical research; and 1 (0.6%) focused on physicians. The screening process is shown in Figure 1.

Finally, we included 42 of 173 (24.2%) English papers in the study; see Table 1. From the perspective of literature research methods, most of the studies were carried out in the form of questionnaires and interviews, and there were 11 (26%) studies in which research models and hypotheses were first proposed and questionnaires were designed for verification to study the mechanism of OHI seeking affecting the physician-patient relationship from different perspectives [11,23-32].

Figure 1. Flowchart of the selection process (English papers). OHI: online health information.

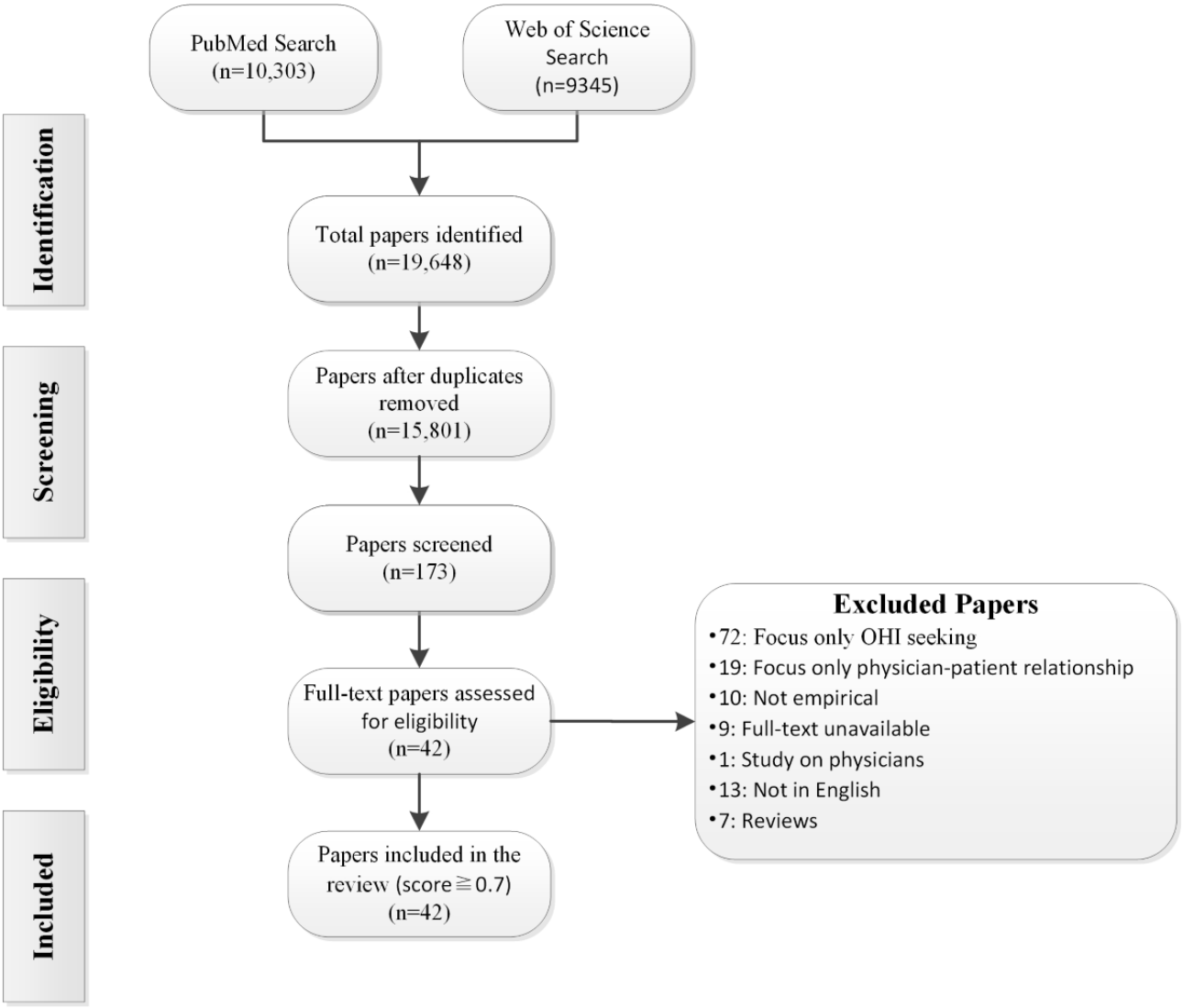


Table 1. Summary of included English papers (n=42).

Number	Reference	Country	Method	Participants, n	Participant characteristics	Factors affecting OHI[a] use covered in this paper (Y=yes/N=no)	Impact of OHI seeking on patients covered in this paper (Y=yes/N=no)	OHI seeking affects patients' evaluation of physicians covered in this paper (Y=yes/N=no)	Discussion of OHI and physician-patient relationship covered in this paper (Y=yes/N=no)
1	[33]	United Kingdom	Semistructured interview	22 (12 [55%] female, 10 [45%] male)	Adult patients with psychosis	Y	Y	N	Y
2	[7]	Austria	Email interview	562 (332 [59.1%] female, 230 [40.9%] male)	Internet citizens	N	N	N	Y
3	[34]	United Kingdom	Survey	202 (102 [50.5%] female, 100 [49.5%] male)	Consecutive adult hematology clinic patients	Y	Y	Y	N
4	[35]	Austria	Semistructured interview	26 (12 [46%] female, 14 [54%] male)	Patients with schizophrenia	N	N	Y	Y
5	[36]	United States	Survey	154 (98 [63.6%] female, 48 [31.2%] male, 8 [5.2%] missing data)	Patients at 3 osteopathic primary care medical clinics	N	Y	N	Y
6	[37]	United Kingdom	Focus group interview	34 (12 [35%] female, 22 [65%] male)	Adult patients with diabetes mellitus, ischemic heart disease, or hepatitis C	N	N	N	Y
7	[38]	United Kingdom	Email interview	31 (28 [90%] female, 3 [10%] male)	Health information seekers	N	Y	N	Y
8	[39]	Saudi Arabia	Survey	431 (181 [41.9%] female, 250 [58.1%] male)	Adult dermatology outpatients	Y	N	N	Y
9	[40]	Canada	Semistructured interview	56 (30 [54%] female, 26 [46%] male)	Adults aged ≥50 years	N	Y	N	Y
10	[41]	United States	Survey and semistructured interview	120 (92 [76.6%] female, 28 [23.3%] male)	Patients new to the rheumatology clinic	N	N	N	Y
11	[42]	United States	Survey	70 (42 [60%] recent internet users [RIUs], 28 [40%] ever internet users [EIUs])	Breast cancer patients	Y	N	N	Y
12	[43]	United States	Survey and semistructured interview	61 (49 [80%] female, 12 [20%] male)	New patients with multiple sclerosis	N	N	N	Y
13	[25]	United States	Individual and focus group interview	20 (11 [55%] female, 9 [45%] male)	Older adults	N	Y	N	N

Number	Reference	Country	Method	Participants, n	Participant characteristics	Factors affecting OHI[a] use covered in this paper (Y=yes/N=no)	Impact of OHI seeking on patients covered in this paper (Y=yes/N=no)	OHI seeking affects patients' evaluation of physicians covered in this paper (Y=yes/N=no)	Discussion of OHI and physician-patient relationship covered in this paper (Y=yes/N=no)
14	[44]	United Kingdom	Telephone survey	3209 (1765 [55%] female, 1444 [45%] male)	A household probability sample from the 48 contiguous states	N	Y	N	Y
15	[45]	United Kingdom	Semistructured interview	47 (32 [68%] female, 15 [32%] male)	Patients with contact with health services for information/treatment in relation to hormone replacement therapy (HRT)/menopause and Viagra/erectile dysfunction	N	N	N	Y
16	[46]	Australia	Survey	93 (44 [47%] female, 49 [53%] male)	Oncology patients	N	Y	N	Y
17	[47]	United Kingdom	Telephone survey	15 females (8 [53%] high school certificate, 5 [33%] bachelor's degree, 2 [14%] postgraduate degree)	Women faced with decisions concerning menopause and HRT	N	Y	N	N
18	[48]	Italy	Survey	1039 (704 [67.76%] female, 335 [32.24%] male)	Adults aged ≥18 years selected from among parents of public school students	Y	Y	N	N
19	[49]	United States	Survey	5075 (3141 [61.89%] female, 1934 [38.11%] male)	Participants in the Health Information National Trends Survey 2007	Y	N	N	N
20	[50]	Australia	Survey	400 (192 [48%] female, 208 [52%] male)	Adult emergency department patients	Y	Y	N	N
21	[51]	United States	Structured in-person interview	1142 (346 [30.29%] female, 796 [69.71%] male)	Adults hospitalized for acute coronary syndromes	Y	N	N	Y
22	[9]	Australia	Interview	33 males	Patients with prostate cancer	N	Y	N	Y
23	[10]	Switzerland	Semistructured interview	32 patients (12 [38%] female, 20 [62%] male) and 20 physicians (4 [20%] female, 16 [80%] male)	Patients and physicians from primary care and medical specialist practices	N	Y	N	Y
24	[52]	Israel	Survey	138 (83 [50.7%] female, 54 [39.3%] male)	Patients at 10 primary care clinics	Y	Y	N	N

Number	Reference	Country	Method	Participants, n	Participant characteristics	Factors affecting OHI[a] use covered in this paper (Y=yes/N=no)	Impact of OHI seeking on patients covered in this paper (Y=yes/N=no)	OHI seeking affects patients' evaluation of physicians covered in this paper (Y=yes/N=no)	Discussion of OHI and physician-patient relationship covered in this paper (Y=yes/N=no)
25	[53]	United States	Telephone survey	2010 (1214 [60.39%] female, 796 [39.61%] male)	Participants in the Surveying the Digital Future, Year 4, national survey	N	N	Y	N
26	[54]	Canada	Survey	39 (27 [70%] female, 11 [28%] male, 1 [2%] unknown)	Patients with thyroid cancer attending appointments with radiation oncologists at 2 tertiary cancer centers	N	Y	N	Y
27	[12]	Switzerland	Questionnaire survey	459 (207 [45.1%] female, 252 [54.9%] male)	460 patients aged ≥18 years	Y	N	N	Y
28	[26]	Netherlands	Questionnaire survey，interview	90 (31 [34%] female, 59 [66%] male)	Patients recently diagnosed with colorectal cancer recruited from 6 hospitals in the Netherlands	Y	Y	N	N
29	[55]	United States	Questionnaire survey	30 (15 [50%] female, 15 [50%] male)	Consecutive patients presenting for preoperative consults for hernia repair requiring surgical mesh	Y	Y	N	Y
30	[56]	Romania	Questionnaire survey	485 (242 [49.9%] female, mean age 50.42 years)	Adult patients	Y	Y	N	N
31	[11]	Singapore	Web-based questionnaire survey	423 (209 [49.4%] female, 214 [50.6%] male)	Internet users	Y	N	N	Y
32	[57]	Malaysia	Questionnaire survey	381 (239 [62.7%] female, 142 [37.3%] male)	Patients in a hospital-based primary care clinic in the University of Malaya Medical Centre	Y	N	N	Y
33	[58]	Belgium	Qualitative semistructured interview	40 (22 [55%] female, 18 [45%] male)	Adults between the ages of 50 and 64 years (middle-aged adults) and 65 and 80 years (older adults)	Y	N	N	Y
34	[59]	China	Focus group interview	46 (34 [74%] cancer patients, 12 [26%] family members)	Patients with cancer or their families	Y	Y	N	Y

Number	Reference	Country	Method	Participants, n	Participant characteristics	Factors affecting OHI[a] use covered in this paper (Y=yes/N=no)	Impact of OHI seeking on patients covered in this paper (Y=yes/N=no)	OHI seeking affects patients' evaluation of physicians covered in this paper (Y=yes/N=no)	Discussion of OHI and physician-patient relationship covered in this paper (Y=yes/N=no)
35	[23]	China	Survey	668 (320 [47.9%] preuse internet samples, 348 [52.1%] not-use internet samples)	Internet citizens	N	Y	N	N
36	[24]	China	Survey	336 (180 [53.6%] female, 156 [46.4%] male)	Participants who underwent treatment with a month	N	Y	Y	N
37	[27]	China	Web-based questionnaire survey	336 (180 [53.6%] female, 156 [46.4%] male)	Chinese individuals who received treatment in the past month and searched the internet for health information	N	Y	Y	N
38	[28]	China	Questionnaire survey	316 (194 [61.4%] female, 122 [38.6%] male)	OHC[b] users	N	Y	Y	Y
39	[29]	China	Web-based questionnaire survey	280 (114 [40.7%] female, 166 [59.3%] male)	Patients who visited the hospital within the past half year or who are visiting the doctor for the first time	N	Y	Y	N
40	[60]	Hong Kong	Questionnaire survey	1179 (717 [60.81%] female, 462 [39.19%] male)	Patients attending the primary care clinic of a university in Hong Kong	Y	Y	N	Y
41	[31]	China	Questionnaire survey	446 (224 [50.2%] female, 222 [49.8%] male)	Patients in Tongji Hospital in Wuhan and the Huazhong University of Science and Technology hospital	N	Y	Y	N
42	[30]	China	Online survey	336 (180 [53.6%] female, 156 [46.4%] male)	Chinese individuals who have experience seeking health information and going to hospitals within the previous month	Y	N	N	N

[a]OHI: online health information.

[b]OHC: online health community.

With searching in the CNKI and SinoMed, there were 5440 initially searched papers, of which 5219 (95.94%) exported papers in Chinese remained after duplication removal using Endnote 20. Of these, 88 (1.69%) papers were included according to the title and abstract for further screening. Through the application of inclusion and exclusion criteria, of these 88 papers, 54 (61%) did not involve the impact of the physician-patient relationship; 10 (11%) did not have health information seeking as the main research object; 7 (8%) were on nonempirical research; for 3 (3%), the original text could not be obtained; and 1 (1%) focused on physicians. The screening process is shown in Figure 2. All 13 (15%) papers met the quality rating except for 2 (15%). Finally, 11 of 88 (13%) Chinese papers were included in this study; see Table 2.

The conclusions of these papers were divided into 5 themes: (1) factors that affect patients' use of OHI, (2) the impact of OHI on patients, (3) OHI seeking affecting patients' evaluation of physicians, (4) discussion with physicians about OHI affecting the physician-patient relationship, and (5) the impact of OHI seeking on the physician-patient relationship, including positive effects, negative effects, and neutral views (Table 3). The neutral view refers to no significant effect or both positive and negative effects.

Figure 2. Flowchart of the selection process (Chinese papers). OHI: online health information.

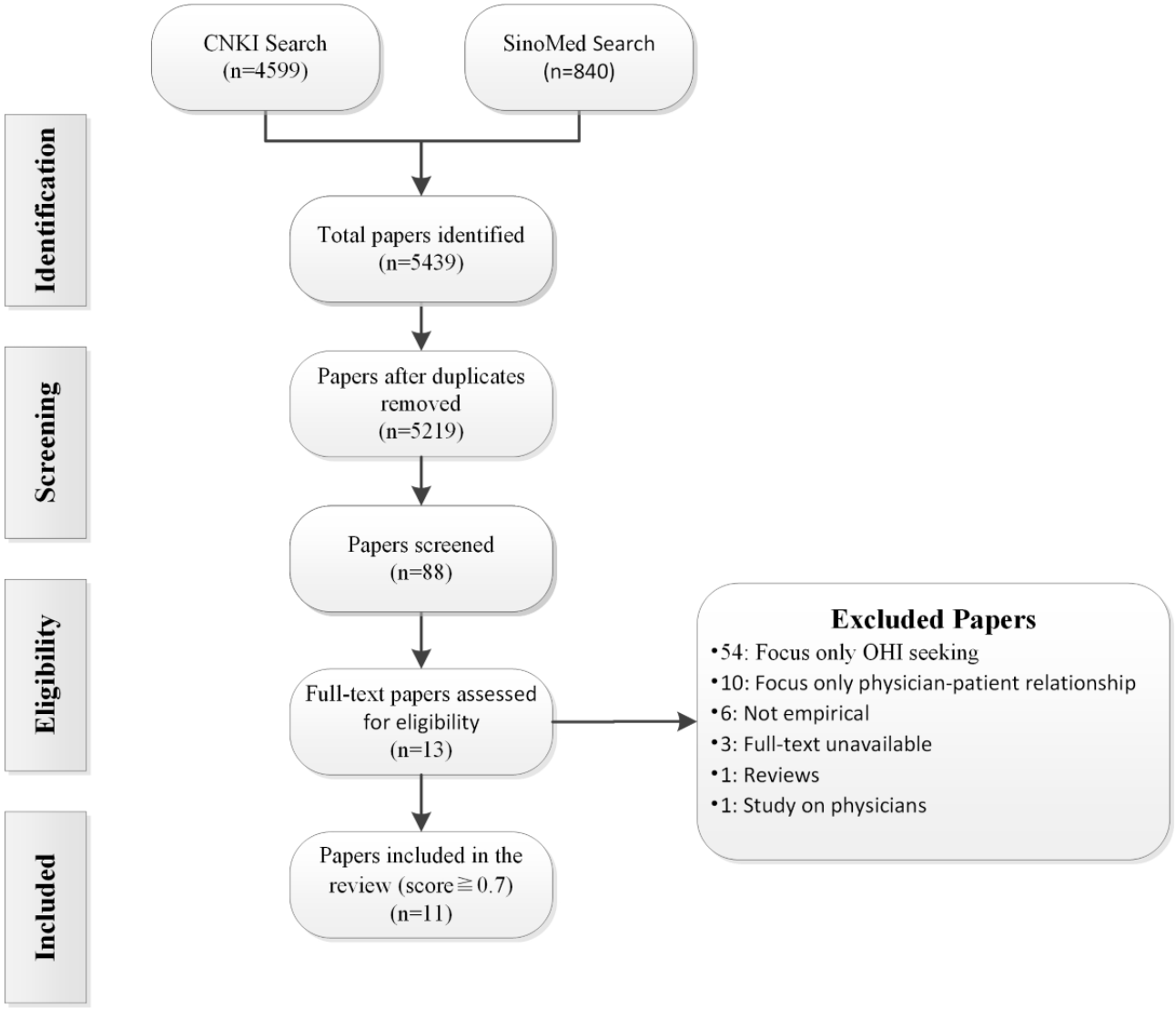


Table 2. Summary of included Chinese papers (n=11).

Number	Reference	Country	Method	Participants, n	Participant characteristics	Factors affecting OHI[a] use covered in this paper (Y=yes/N=no)	Impact of OHI seeking on patients covered in this paper (Y=yes/N=no)	OHI seeking affects patients' evaluation of physicians covered in this paper (Y=yes/N=no)	Discussion of OHI and physician-patient relationship covered in this paper (Y=yes/N=no)
1	[61]	China	Survey	179 (85 [47.5%] female, 94 [52.5%] male）	Outpatients with chronic diseases	N	N	N	Y
2	[62]	China	Survey, interview	467 (277 [59.3%] female, 190 [40.7%] male)	Chinese citizens	Y	N	N	N
3	[63]	China	Survey	446 (224 [50.2%] female, 222 [49.8%] male)	Health information seekers	N	Y	N	N
4	[64]	China	Survey	951 patients (495 [52.1%] female, 456 [47.9%] male) and 888 physicians (348 [39.2%] female, 540 [60.8%] male)	Patients over 18 and doctors in each department	Y	N	Y	N
5	[65]	China	Survey	1232 (611 [49.59%] users of OHI, 621 [50.41%] nonusers of OHI)	Chinese netizens	N	Y	Y	Y
6	[66]	China	China Family Panel Studies (CFPS) data	29,647 (14,815 [49.97%] female, 14,832 [50.03%] male)	Chinese citizens	Y	N	Y	N
7	[67]	China	2013 Chinese Social Survey data	10,206 (2073 [20.31%] netizens, 4654 [45.6%] nonnetizens, and 3479 [34.09%] missing data)	Chinese citizens	Y	N	Y	N
8	[68]	China	Survey	336 (180 [53.6%] female, 156 [46.4%] male)	ChunYu Doctors website users	N	N	Y	N
9	[69]	China	2018 CFPS adult questionnaire data	25,015 (13,083 [52.3%] female, 11,932 [47.7%] male)	Chinese citizens	Y	N	N	N
10	[70]	China	2011 and 2012 Chinese General Social Survey (CGSS)	5546 (in 2021) and 5797 (in 2012)	Chinese citizens	Y	N	Y	N

Number	Reference	Country	Method	Participants, n	Participant characteristics	Factors affecting OHI[a] use covered in this paper (Y=yes/N=no)	Impact of OHI seeking on patients covered in this paper (Y=yes/N=no)	OHI seeking affects patients' evaluation of physicians covered in this paper (Y=yes/N=no)	Discussion of OHI and physician-patient relationship covered in this paper (Y=yes/N=no)
11	[32]	China	Questionnaire survey	464 (241 [51.9%] female, 253 [48.1%] male)	Chinese citizens	Y	Y	Y	N

[a]OHI: online health information.

Table 3. OHI[a] seeking affects physician-patient relationships.

Number	Reference	Country	Impact of OHI seeking on physician-patient relationship		
			Positive effects covered in this paper (Y=yes/N=no)	Neutral views covered in this paper (Y=yes/N=no)	Negative effects covered in this paper (Y=yes/N=no)
1	[33]	United Kingdom	Y	N	N
2	[7]	Austria	N	Y	N
3	[34]	United Kingdom	Y	N	N
4	[35]	Austria	Y	N	N
5	[36]	United States	Y	N	N
6	[37]	United Kingdom	Y	N	N
7	[38]	United Kingdom	Y	N	N
8	[39]	Saudi Arabia	Y	N	N
9	[40]	Canada	N	Y	N
10	[41]	United States	N	Y	N
11	[42]	United States	N	Y	N
12	[43]	United States	N	Y	N
13	[25]	United States	Y	N	N
14	[44]	United Kingdom	Y	N	N
15	[45]	United Kingdom	N	Y	N
16	[46]	Australia	Y	N	N
17	[47]	United Kingdom	Y	N	N
18	[48]	Italy	Y	N	N
19	[49]	United States	N	Y	N
20	[50]	Australia	Y	N	N
21	[51]	United States	Y	N	N
22	[9]	Australia	N	Y	N
23	[10]	Switzerland	N	Y	N
24	[52]	Israel	Y	N	N
25	[53]	United States	Y	N	N
26	[54]	Canada	N	Y	N
27	[12]	Switzerland	N	Y	N
28	[26]	Netherlands	N	Y	N
29	[55]	United States	N	N	Y
30	[56]	Romania	Y	N	N
31	[11]	Singapore	Y	N	N
32	[57]	Malaysia	N	N	Y
33	[58]	Belgium	Y	N	N
34	[59]	China	Y	N	N
35	[23]	China	N	Y	N
36	[24]	China	Y	N	N
37	[27]	China	Y	N	N
38	[28]	China	Y	N	N
39	[29]	China	Y	N	N
40	[60]	China	N	N	Y

Number	Reference	Country	Impact of OHI seeking on physician-patient relationship		
			Positive effects covered in this paper (Y=yes/N=no)	Neutral views covered in this paper (Y=yes/N=no)	Negative effects covered in this paper (Y=yes/N=no)
41	[31]	China	Y	N	N
42	[30]	China	Y	N	N
43	[61]	China	Y	N	N
44	[62]	China	N	N	Y
45	[63]	China	N	Y	N
46	[64]	China	Y	N	N
47	[65]	China	Y	N	N
48	[66]	China	N	N	Y
49	[67]	China	N	N	Y
50	[68]	China	Y	N	N
51	[69]	China	N	N	Y
52	[70]	China	N	N	Y
53	[32]	China	Y	N	N

[a]OHI: online health information.

Factors Affecting the Use of OHI

Studies show that education level, income, gender, age, health literacy, culture, and other factors can affect people's use of OHI [12,26,33,34,39,42,49,51,56]. Five papers showed that users with high education level and high income are more willing to use OHI [12,33,34,39,42,62]. This population has relatively high health literacy and can better deal with OHI. Patients with difficulties in understanding health information are less likely to ask questions or seek guidance during consultation. Gantenbein et al [12] found that women are more willing to conduct OHI searches, while Aref-Adib et al [33] found that young male psychiatric patients are more likely to discuss health information with their physicians. De Looper et al [26] and Drug et al [56] found that younger patients engage more in OHIS, but Waring et al [51] did not observe the age difference possibly because the large age-grouping scope could not reflect the difference between the elderly and the young. In addition to personal factors, Chiu et al [59] found that the cultural environment of patients may also affect the communication on health factors. In a hierarchical culture of patients and physicians, patients are unwilling to ask questions for fear that the physicians would be unhappy. Instead, they choose to listen to the advice of physicians [59]. In this paper, we only included papers studying the relationship between physicians and patients; however, maybe many other factors also affect the use of OHI by patients.

Impact of OHI Seeking on Patients

Several studies have shown that OHI can enhance the communication ability and decision-making ability of patients. The study conducted by Iverson et al [36] showed that 46% of patients said they would change their health-related behaviors after searching for health information online. After searching the health information online, patients have a certain understanding of their own health status and disease treatment and can better understand the medical terms used by physicians when talking with them [25,38,59]. Murray et al [44] showed that people who discuss health information with physicians often have higher self-assessment ability to assess their own health. Liang et al [61] showed that patients who think that OHI is important and helpful to health decision making are more inclined to think that it will be beneficial to the physician-patient relationship.

However, some patients may show some negative effects after OHI query. Aref-Adib et al [33] found that some patients may have concerns over what they read and that they change medication adherence and behavior without communication with the physicians. Another concern raised by OHI seeking is related to the quality of OHI, such as the credibility and limitations of information [40]. OHI will affect patients' decision making, but patients still regard physicians as the main source of health information [47]. Due to the uneven quality of OHI and the lack of quality control, in addition to patients lacking medical information literacy, the judgments made by patients based on OHI are generally unscientific and difficult to be recognized by doctors, which may have a negative impact on the physician-patient relationship [62].

OHI Seeking Affects Patients' Evaluation of Physicians

The impact of OHI is mainly reflected in patients' trust in and satisfaction with physicians [42]. Patients' satisfaction with physicians is composed of many factors, among which the the main influencing factors are related to the actual communication between patients and physicians [23]. However, most patients are afraid to challenge their doctors, so they are reluctant to discuss their OHI [41]. Patients' satisfaction with OHI has a direct and positive impact on psychological safety [34], while psychological safety might have a direct and positive impact on patients' trust in physicians [24]. When patients use the network health community, the trust relationship among

community members also affects patients' trust in and satisfaction with physicians [71]. Liu [68] showed that continuous use of online health communities (OHCs) increases users' satisfaction with medical services.

Discussion With Physicians About Health Information Affects Physician-Patient Relationships

After OHI seeking, some patients choose to share health information with their physicians. Part of the motivation for discussing health information with physicians is that patients want to meet their psychological and emotional needs [42]. Iverson et al [36] showed that 73% of patients like to discuss OHI with physicians mainly because they think physicians are willing to discuss OHI. The willingness of physicians to discuss health information with patients is crucial. After discussing their concerns about OHI with physicians, the patients' medication adherence and behavior remain unchanged and the anxiety caused by OHI reduces [33]. In addition, after discussing OHI with physicians, patients' satisfaction with physicians significantly improves [39]. People who discuss OHI with their physicians think it has a positive impact on the disease and their relationship with their physicians [7].

However, some studies have also proved that the main reason patients do not actively discuss health information with physicians is the fear of challenging physicians' authority [33,35,38,40,42,43,59]. Patients worry that when they talk about OHI, or express some opinions that physicians cannot refute, the physicians will feel criticized [35]. Of course, there are also some health care professionals trying to maintain the existing authority by not discussing OHI [45]. Guanghua [62] showed that the negative impact of OHI is greater, which creates greater obstacles in the communication between doctors and patients in China.

OHI Seeking Affects Physician-Patient Relationships in China

A total of 53 papers were included in this study, of which 31 (58%) hold that OHI seeking has a positive impact on the physician-patient relationship, 14 (26%) have a neutral view, and 8 (15%) have a negative influence. It is worthy of in-depth study that 6 (11%) papers showed that OHI seeking has a negative impact on the physician-patient relationship in China. Therefore, the negative impact of OHI on China's internet users is worthy of attention. Due to the large population of China, the time for each patient to communicate with the doctor is short, and patients choose to search online for health information more for convenience than for accuracy or authority [67]. Some studies have shown that the inclusion of some wrong medical information and reports of malignant incidents in the physician-patient relationship have a negative impact on physician-patient trust, confirming media depression theory [67,69,70]. Feifei [64] found that for ordinary patients, due to the professional barriers of medical knowledge, it is difficult for them to distinguish between true and false after receiving false health information on the internet. This causes patients to question doctors and leads to difficulties in the physician-patient relationship [64]. Therefore, it is important to improve patients' health information literacy and the quality of OHI.

Discussion

Principal Findings

Based on the review of the included studies, we found that there are many factors that affect patients' choice of OHI, such as gender, age, education level, income, health literacy, and culture [33,34,39,42,49,51]. People with high income, education level, and health literacy are more likely to use OHI. The age difference is mainly between the young and the old. There is a digital divide between the elderly and the young [72]. One study found that older people prefer to choose people as sources of information, such as health care providers, pharmacists, relatives, and retired community workers [73]. From the perspective of patients, most of them think that OHI seeking does not affect the physician-patient relationship; some patients think it has a positive impact on the physician-patient relationship, and a few patients think it may have a negative impact on the physician-patient relationship [46]. As the impact of OHI seeking on the physician-patient relationship may be restricted by social and cultural factors, it may have adverse effects in a culture with distinct levels of patients and physicians [59].

The mechanism of OHI seeking affecting the physician-patient relationship is relatively complex. According to the study findings, OHI seeking can enhance patients' understanding of medical knowledge and enhance their decision-making ability and communication ability with physicians. At the same time, OHI seeking can also have an impact on patients' own psychology. Good quality of health information has a positive impact on the psychological safety of patients. Bylund [74] found that high satisfaction with OHI can promote patients' psychological security when communicating with physicians. Psychological safety has a certain impact on the distrust in the physician-patient relationship so as to affect the relationship [75]. Side effects of drugs and other information cause anxiety in patients. As the internet provides an opportunity to communicate with others about their concerns, anxiety tends to increase [76]. A small number of patients even have drug compliance changes and changes in their own medical behavior [76]. Previous studies have shown that that OHI can affect the consistency of communication between physicians and patients and the compliance of patients [33].

It is important that patients discuss health information with their physicians. Patients will seek OHI to prepare for seeing a doctor, fully participate in the decision making, and actively supplement their information during the process of seeking medical service [77]. However, this will also cause anxiety and a series of changes, such as compliance change and medical behavior change [33]. If patients do not discuss health information with physicians, the negative effects on some patients might even worsen. If patients discuss health information with physicians, these negative effects can be eliminated and alleviated. The survey results show that discussing health information with physicians is beneficial to patients' satisfaction with and trust in physicians. We must admit that patients need to discuss their health information with physicians to better promote the physician-patient relationship and improve medical services

[78]. Good physician-patient communication can improve the clinical outcomes of some diseases [79].

Several studies have mentioned that patients are afraid to discuss health information with physicians because they are afraid of challenging the authority of physicians and even of conflicts with physicians. The OHC is not well received by the professional medical staff. They have doubts about the quality of a lot of OHI and whether they can explain the medical information to the patients in a better way [16]. Patients tend to remain silent if they do not feel the physician's willingness to discuss OHI with them. When patients consult about traditional and nontraditional therapies, many physicians react defensively, resulting in adverse effects on patients' trust in them and the communication between physicians and patients [80]. Some physicians try to maintain their authority as physicians by avoiding discussing OHI [45]. Due to the widespread popularity of OHI, physicians should be aware that many patients seek OHI before consultation, and actively discuss and exchange OHI with patients [81].

Of the included 53 papers, 21 (39.6%) studied the impact of OHI seeking on the physician-patient relationship in China, of which 2 (9%) papers specifically mentioned that cultural factors play a potential role in OHI seeking for physician-patient relationships [59,62]. In the culture of hierarchical physician-patient relationships in China, the patient fully follows the physician's recommendations [59]. The popularity of OHI allows patients to play a more important role in the medical process. However, under the medical environment of "more patients, fewer physicians" in China, the communication time between each physician and patient is too short [62]. If the doctor cannot convince the patient and deny the patient's opinion directly without explanation, the conflict weakens the authority of the doctor and exacerbates the negative impact on the physician-patient relationship [62].

Several papers have shown that the internet usage time could reduce the patient's trust in the doctor [66,67]. Medical corruption, medical malpractice, physician-patient conflict, and other contents are more likely to spread among Chinese patients. Various types of negative news is frequently pushed to patients. The negative factors in the physician-patient relationship are magnified. The media often blame medical disputes on medical personnel, which exacerbates patients' distrust of doctors. In addition to negative news, the low quality of OHI has a negative impact on the physician-patient relationship [61,82].

OHI is a double-edged sword for the relationship between physicians and patients. It is becoming increasingly important in the relationship between physicians and patients. With high-quality OHI, it is relatively easier to have a positive impact on patients, thus promoting the physician-patient relationship. With the rapid growth and wide use of medical websites, there are important problems about the necessity of quality control [83]. The pattern of patients' access to health information is changing from passive recipients to active service seekers [77]. Health care professionals should not only discuss health information with patients but also guide them to correctly seek and use health information. Patients who can reasonably understand OHI can reduce the burden of physicians in the consultation and improve the communication [63].

Limitations

This study has a wide range of retrieval. When references were included, the focus was on whether health information seeking has an impact on the relationship between physicians and patients. Papers studying the impact of health information seeking on patients were not included, which may have led to missing potential research. In addition, due to the lack of a large number of studies and more reliable evidence, we could not reach a strong conclusion about how health information seeking affects the physician-patient relationship.

Conclusion

This study mainly focused on the effects of OHI on the relationship between physicians and patients. There are many factors influencing people's use of OHI, and young, female, highly educated, and high-income patients are more willing to search OHI. OHI seeking can affect patients' mentality and behavior. Through understanding OHI, patients can have a better understanding of medical knowledge, improve self-confidence during communication, and enhance self-decision-making behaviors. However, some OHI can lead to negative emotions and even change patients' health behaviors, due to the uneven quality of OHI. OHI seeking also affect patients' evaluation of doctors, including patients' trust in and satisfaction with physicians. OHI users choose to discuss OHI with doctors, which is beneficial to the physician-patient relationship in most cases. However, due to the subjective consciousness of patients, they may be concerned that it might affect the authority of physicians, which is the reason some patients do not initiate the discussion of health information. Moreover, the negative impact of OHI on China's internet users is worthy of attention; due to the low quality of OHI, poor health information literacy, short physician-patient communication time, and various types of negative news, patients' trust in doctors has declined. At present, China's vigorous promotion of "internet + medical health" and the reform of the hierarchical medical will be of great significance to improving the physician-patient communication model and promoting harmonious physician-patient relationships. At the same time, improving people's health information literacy and the quality of OHI is the crucial step in facilitating the positive effects of OHI on the physician-patient relationship.

Acknowledgments

The authors thank the State Key Program of the National Social Science of China (grant no. 17AZD037) and the Natural Science Foundation of Hunan Province (2019JJ80032) for their assistance in this study. We also appreciate the support of all participants.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA 2020 Checklist.
[PDF File (Adobe PDF File), 141 KB - jmir_v24i2e23354_app1.pdf]

Multimedia Appendix 2
CASP (Critical Appraisal Skills Program) quality assessment for qualitative studies.
[PDF File (Adobe PDF File), 190 KB - jmir_v24i2e23354_app2.pdf]

Multimedia Appendix 3
Quality assessment tool for quantitative studies.
[PDF File (Adobe PDF File), 224 KB - jmir_v24i2e23354_app3.pdf]

References

1. Statista. Global Digital Population as of January 2021. URL: https://www.statista.com/statistics/617136/digital-population-worldwide/ [accessed 2021-10-15]
2. Berger M, Wagner TH, Baker LC. Internet use and stigmatized illness. Soc Sci Med 2005 Oct;61(8):1821-1827. [doi: 10.1016/j.socscimed.2005.03.025] [Medline: 16029778]
3. Bhandari N, Shi Y, Jung K. Seeking health information online: does limited healthcare access matter? J Am Med Inform Assoc 2014;21(6):1113-1117 [FREE Full text] [doi: 10.1136/amiajnl-2013-002350] [Medline: 24948558]
4. Valero-Aguilera B, Bermúdez-Tamayo C, García-Gutiérrez JF, Jiménez-Pernett JH, Cózar-Olmo JM, Guerrero-Tejada R, et al. Information needs and internet use in urological and breast cancer patients. Support Care Cancer 2014 Feb;22(2):545-552. [doi: 10.1007/s00520-013-2009-y] [Medline: 24122406]
5. Gray NJ, Klein JD, Noyce PR, Sesselberg TS, Cantrill JA. Health information-seeking behaviour in adolescence: the place of the internet. Soc Sci Med 2005 Apr;60(7):1467-1478. [doi: 10.1016/j.socscimed.2004.08.010] [Medline: 15652680]
6. Lemire M, Paré G, Sicotte C, Harvey C. Determinants of internet use as a preferred source of information on personal health. Int J Med Inform 2008 Nov;77(11):723-734. [doi: 10.1016/j.ijmedinf.2008.03.002] [Medline: 18434246]
7. Haluza D, Naszay M, Stockinger A, Jungwirth D. Digital natives versus digital immigrants: influence of online health information seeking on the doctor-patient relationship. Health Commun 2017 Nov;32(11):1342-1349. [doi: 10.1080/10410236.2016.1220044] [Medline: 27710132]
8. Pal K, Eastwood SV, Michie S, Farmer AJ, Barnard ML, Peacock R, et al. Computer-based diabetes self-management interventions for adults with type 2 diabetes mellitus. Cochrane Database Syst Rev 2013 Mar 28(3):CD008776 [FREE Full text] [doi: 10.1002/14651858.CD008776.pub2] [Medline: 23543567]
9. Broom A. Virtually he@lthy: the impact of internet use on disease experience and the doctor-patient relationship. Qual Health Res 2005 Mar;15(3):325-345. [doi: 10.1177/1049732304272916] [Medline: 15761103]
10. Sommerhalder K, Abraham A, Zufferey MC, Barth J, Abel T. Internet information and medical consultations: experiences from patients' and physicians' perspectives. Patient Educ Couns 2009 Nov;77(2):266-271. [doi: 10.1016/j.pec.2009.03.028] [Medline: 19411157]
11. Jiang S, Liu J. Examining the relationship between internet health information seeking and patient-centered communication in China: taking into account self-efficacy in medical decision-making. Chin J Commun 2020 May 25;13(4):407-424. [doi: 10.1080/17544750.2020.1769700]
12. Gantenbein L, Navarini AA, Maul LV, Brandt O, Mueller SM. Internet and social media use in dermatology patients: search behavior and impact on patient-physician relationship. Dermatol Ther 2020 Nov;33(6):e14098. [doi: 10.1111/dth.14098] [Medline: 32725746]
13. Erdem SA, Harrison-Walker LJ. The role of the internet in physician–patient relationships: the issue of trust. Business Horizons 2006 Sep;49(5):387-393. [doi: 10.1016/j.bushor.2006.01.003]
14. Georgopoulou S, Prothero L, D'Cruz DP. Physician-patient communication in rheumatology: a systematic review. Rheumatol Int 2018 May 26;38(5):763-775 [FREE Full text] [doi: 10.1007/s00296-018-4016-2] [Medline: 29582095]
15. Kaba R, Sooriakumaran P. The evolution of the doctor-patient relationship. Int J Surg 2007 Feb;5(1):57-65 [FREE Full text] [doi: 10.1016/j.ijsu.2006.01.005] [Medline: 17386916]
16. Wald HS, Dube CE, Anthony DC. Untangling the web: the impact of internet use on health care and the physician-patient relationship. Patient Educ Couns 2007 Nov;68(3):218-224. [doi: 10.1016/j.pec.2007.05.016] [Medline: 17920226]
17. Emanuel EJ, Emanuel LL. Four models of the physician-patient relationship. JAMA 1992 Apr 22;267(16):2221-2226. [doi: 10.1001/jama.267.16.2221] [Medline: 1556799]
18. National Bureau of Statistics. Bulletin of the Seventh National Census (No.2). 2021 May 11. URL: http://www.stats.gov.cn/ztjc/zdtjgz/zgrkpc/dqcrkpc/ggl/202105/t20210519_1817695.html [accessed 2021-10-15]

19. Hong-bin X, Shou-huo J, Wen B, Yi-ming X, Xin-yi S, Zuo-fan X, et al. Integration of internet health and "Healthy China" strategy. Shanghai J Prev Med 2021;33(08):659-663. [doi: 10.19428/j.cnki.sjpm.2021.20931]
20. Si-wen L, Zhao-xia G, Shi-jun L, Xin-yi M, Bao-jia Z, Jun-xue W. Analysis on the significance and necessity of "Internet +" medical mode. China Market 2021;28(20):187-188. [doi: 10.13939/j.cnki.zgsc.2021.20.187]
21. Critical Appraisal Skills Programme. CASP Checklist: 10 Questions to Help You Make Sense of a Qualitative Research. 2018. URL: https://casp-uk.b-cdn.net/wp-content/uploads/2018/03/CASP-Qualitative-Checklist-2018_fillable_form.pdf [accessed 2021-10-15]
22. Tan SS, Goonawardene N. Internet health information seeking and the patient-physician relationship: a systematic review. J Med Internet Res 2017 Jan 19;19(1):e9 [FREE Full text] [doi: 10.2196/jmir.5729] [Medline: 28104579]
23. Tian L, Hong C, Yun-jie X, Cheng-hong Z. Internet usage, physician performances and patient's trust in physician during diagnoses: investigating both pre-use and not-use internet groups. : IEEE; 2016 Presented at: 2016 49th Hawaii International Conference on System Sciences (HICSS); Jan 5-8, 2016; Koloa. [doi: 10.1109/hicss.2016.400]
24. Zhang R, Lu X, Wu W, Shang X, Liu M. Mature or emerging? The impact of treatment-related internet health information seeking on patients' trust in physicians. Int J Environ Res Public Health 2018 Aug 28;15(9):1855 [FREE Full text] [doi: 10.3390/ijerph15091855] [Medline: 30154309]
25. Xie B. Older adults' health information wants in the internet age: implications for patient-provider relationships. J Health Commun 2009 Sep;14(6):510-524. [doi: 10.1080/10810730903089614] [Medline: 19731124]
26. de Looper M, van Weert JCM, Schouten B, Bolle S, Belgers EHJ, Eddes EH, et al. The influence of online health information seeking before a consultation on anxiety, satisfaction, and information recall, mediated by patient participation: field study. J Med Internet Res 2021 Jul 05;23(7):e23670 [FREE Full text] [doi: 10.2196/23670] [Medline: 34255657]
27. Zhang Z, Yang H, He J, Lu X, Zhang R. The impact of treatment-related internet health information seeking on patient compliance. Telemed J E Health 2021 May 01;27(5):513-524. [doi: 10.1089/tmj.2020.0081] [Medline: 32907505]
28. Lu X, Zhang R. Impact of patient information behaviours in online health communities on patient compliance and the mediating role of patients' perceived empathy. Patient Educ Couns 2021 Jan;104(1):186-193. [doi: 10.1016/j.pec.2020.07.001] [Medline: 32665071]
29. Yongmei L, Xiangxi L. Effect of patient online information searching on the trust in the doctor: a cognitive dissonance theory perspective. 2019 Sep 12 Presented at: 2019 Chinese Control And Decision Conference (CCDC); June 3-5, 2019; Nanchang, China p. 4254-4259. [doi: 10.1109/ccdc.2019.8833097]
30. Zhang R, Lu X, Wu WF, Shang X. Why do patients follow physicians' advice? The influence of patients' regulatory focus on adherence: an empirical study in China. BMC Health Serv Res 2019 May 10;19(1):301 [FREE Full text] [doi: 10.1186/s12913-019-4127-9] [Medline: 31077196]
31. Peng Y, Yin PO, Deng Z, Wang R. Patient-physician interaction and trust in online health community: the role of perceived usefulness of health information and services. Int J Environ Res Public Health 2019 Dec 24;17(1):139 [FREE Full text] [doi: 10.3390/ijerph17010139] [Medline: 31878145]
32. Tianyue L. Study on Continuance Use Behavior of Online Healthcare Community and Its Impact on Doctor-Patient Relationship. Hefei: Hefei University of Technology; 2017. URL: https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD201702&filename=1017179124.nh [accessed 2021-10-15]
33. Aref-Adib G, O'Hanlon P, Fullarton K, Morant N, Sommerlad A, Johnson S, et al. A qualitative study of online mental health information seeking behaviour by those with psychosis. BMC Psychiatry 2016 Jul 11;16(1):232 [FREE Full text] [doi: 10.1186/s12888-016-0952-0] [Medline: 27400874]
34. Rider T, Malik M, Chevassut T. Haematology patients and the internet: the use of on-line health information and the impact on the patient-doctor relationship. Patient Educ Couns 2014 Nov;97(2):223-238. [doi: 10.1016/j.pec.2014.06.018] [Medline: 25053474]
35. Schrank B, Sibitz I, Unger A, Amering M. How patients with schizophrenia use the internet: qualitative study. J Med Internet Res 2010 Dec 19;12(5):e70 [FREE Full text] [doi: 10.2196/jmir.1550] [Medline: 21169176]
36. Iverson S, Howard K, Penney B. Impact of internet use on health-related behaviors and the patient-physician relationship: a survey-based study and review. J Am Osteopath Assoc 2008 Dec;108(12):699-711. [doi: 10.7556/jaoa.2008.108.12.699] [Medline: 19075034]
37. Stevenson FA, Kerr C, Murray E, Nazareth I. Information from the internet and the doctor-patient relationship: the patient perspective; a qualitative study. BMC Fam Pract 2007 Aug 16;8:47 [FREE Full text] [doi: 10.1186/1471-2296-8-47] [Medline: 17705836]
38. Kivits J. Informed patients and the internet: a mediated context for consultations with health professionals. J Health Psychol 2006 Mar;11(2):269-282. [doi: 10.1177/1359105306061186] [Medline: 16464924]
39. AlGhamdi KM, Almohideb MA. Internet use by dermatology outpatients to search for health information. Int J Dermatol 2011 Mar;50(3):292-299. [doi: 10.1111/j.1365-4632.2010.04705.x] [Medline: 21342162]
40. Silver MP. Patient perspectives on online health information and communication with doctors: a qualitative study of patients 50 years old and over. J Med Internet Res 2015 Jan 13;17(1):e19 [FREE Full text] [doi: 10.2196/jmir.3588] [Medline: 25586865]

41. Hay MC, Cadigan RJ, Khanna D, Strathmann C, Lieber E, Altman R, et al. Prepared patients: internet information seeking by new rheumatology patients. Arthritis Rheum 2008 Apr 15;59(4):575-582 [FREE Full text] [doi: 10.1002/art.23533] [Medline: 18383399]
42. D'Agostino TA, Ostroff JS, Heerdt A, Dickler M, Li Y, Bylund CL. Toward a greater understanding of breast cancer patients' decisions to discuss cancer-related internet information with their doctors: an exploratory study. Patient Educ Couns 2012 Oct;89(1):109-115 [FREE Full text] [doi: 10.1016/j.pec.2012.05.008] [Medline: 22722063]
43. Hay MC, Strathmann C, Lieber E, Wick K, Giesser B. Why patients go online: multiple sclerosis, the internet, and physician-patient communication. Neurologist 2008 Nov;14(6):374-381. [doi: 10.1097/NRL.0b013e31817709bb] [Medline: 19008743]
44. Murray E, Lo B, Pollack L, Donelan K, Catania J, White M, et al. The impact of health information on the internet on the physician-patient relationship: patient perceptions. Arch Intern Med 2003 Jul 28;163(14):1727-1734. [doi: 10.1001/archinte.163.14.1727] [Medline: 12885689]
45. Hart A, Henwood F, Wyatt S. The role of the Internet in patient-practitioner relationships: findings from a qualitative research study. J Med Internet Res 2004 Sep 30;6(3):e36 [FREE Full text] [doi: 10.2196/jmir.6.3.e36] [Medline: 15471762]
46. Newnham GM, Burns WI, Snyder RD, Dowling AJ, Ranieri NF, Gray EL, et al. Information from the internet: attitudes of Australian oncology patients. Intern Med J 2006 Nov;36(11):718-723. [doi: 10.1111/j.1445-5994.2006.01212.x] [Medline: 17040358]
47. Sillence E, Briggs P, Harris PR, Fishwick L. How do patients evaluate and make use of online health information? Soc Sci Med 2007 May;64(9):1853-1862. [doi: 10.1016/j.socscimed.2007.01.012] [Medline: 17328998]
48. Bianco A, Zucco R, Nobile CGA, Pileggi C, Pavia M. Parents seeking health-related information on the internet: cross-sectional study. J Med Internet Res 2013 Sep 18;15(9):e204 [FREE Full text] [doi: 10.2196/jmir.2752] [Medline: 24047937]
49. Chung JE. Patient-provider discussion of online health information: results from the 2007 Health Information National Trends Survey (HINTS). J Health Commun 2013;18(6):627-648. [doi: 10.1080/10810730.2012.743628] [Medline: 23590202]
50. Cocco AM, Zordan R, Taylor DM, Weiland TJ, Dilley SJ, Kant J, et al. Dr Google in the ED: searching for online health information by adult emergency department patients. Med J Aust 2018 Oct 15;209(8):342-347. [doi: 10.5694/mja17.00889] [Medline: 30107763]
51. Waring ME, McManus DD, Amante DJ, Darling CE, Kiefe CI. Online health information seeking by adults hospitalized for acute coronary syndromes: who looks for information, and who discusses it with healthcare providers? Patient Educ Couns 2018 Nov;101(11):1973-1981 [FREE Full text] [doi: 10.1016/j.pec.2018.06.016] [Medline: 30305253]
52. Russ H, Giveon SM, Catarivas MG, Yaphe J. The effect of the Internet on the patient-doctor relationship from the patient's perspective: a survey from primary care. Isr Med Assoc J 2011 Apr;13(4):220-224 [FREE Full text] [Medline: 21598810]
53. Ybarra M, Suman M. Reasons, assessments and actions taken: sex and age differences in uses of internet health information. Health Educ Res 2008 Jun;23(3):512-521. [doi: 10.1093/her/cyl062] [Medline: 16880222]
54. Chang K, Berthelet E, Grubbs E, Hamilton S, Karvat A, Tran E, et al. Websites, websites everywhere: how thyroid cancer patients use the internet. J Cancer Educ 2020 Dec;35(6):1177-1183. [doi: 10.1007/s13187-019-01576-5] [Medline: 31332622]
55. Miller MP, Arefanian S, Blatnik JA. The impact of internet-based patient self-education of surgical mesh on patient attitudes and healthcare decisions prior to hernia surgery. Surg Endosc 2020 Nov;34(11):5132-5141. [doi: 10.1007/s00464-019-07300-0] [Medline: 31832857]
56. Drug VL, Chirila I, Albusoda A, Bancila I, Iordache S, Nedelcu L, et al. How the internet influences the relationship between outpatients and gastroenterologists: a multicenter study. Turk J Gastroenterol 2020 Jan;31(1):17-22 [FREE Full text] [doi: 10.5152/tjg.2019.18353] [Medline: 31895040]
57. Lim HM, Wong SS, Yip KC, Chang FWS, Chin AJZ, Teo CH, et al. Online health information-seeking behaviour of patients attending a primary care clinic in Malaysia: a cross-sectional study. Fam Pract 2021 Aug 23:cmab099. [doi: 10.1093/fampra/cmab099] [Medline: 34423368]
58. Huisman M, Joye S, Biltereyst D. Searching for health: Doctor Google and the shifting dynamics of the middle-aged and older adult patient-physician relationship and interaction. J Aging Health 2020 Oct;32(9):998-1007. [doi: 10.1177/0898264319873809] [Medline: 31517558]
59. Chiu YC. Probing, impelling, but not offending doctors: the role of the internet as an information source for patients' interactions with doctors. Qual Health Res 2011 Dec;21(12):1658-1666. [doi: 10.1177/1049732311417455] [Medline: 21799204]
60. Wong DK, Cheung MK. Online health information seeking and eHealth literacy among patients attending a primary care clinic in Hong Kong: a cross-sectional survey. J Med Internet Res 2019 Mar 27;21(3):e10831 [FREE Full text] [doi: 10.2196/10831] [Medline: 30916666]
61. Hui L, Ye Z, Haiyun Q, Ying Z, Xia Q. The impact of network health information search on the relationship between doctors and patients among patients with chronic diseases. Chin Rural Health Serv Admin 2019;39(3):199-202. [doi: 10.3969/j.issn.1005-5916.2019.03.012]
62. Guanghua W. The Influence of Online Health Information Seeking Behavior on Doctor-Patient Communication. Nanjing: Nanjing University; 2018. URL: https://tinyurl.com/2p8nvaah [accessed 2021-10-15]

63. Pingping Y, Zhaohua D, Xicheng L. Study on the influence of online health information seeking on doctor-patient relationship. J Med Intell 2017;38(10):2-7,35 [FREE Full text] [doi: 10.3969/j.issn.1673-6036.2017.10.001]
64. Feifei D. Research on the Doctor-Patient Relationship under the Environment of Network. Canton: Southern Medical University; 2014. URL: https://tinyurl.com/2p8w7n98 [accessed 2021-10-15]
65. Zhijie X, Bingzhang Z, Deshuai Z, Youdong Z. The impact of the use of online health resources on the doctor-patient relationship. China Health Ind 2015;12(29):1-4. [doi: 10.16659/j.cnki.1672-5654.2015.29.001]
66. Xinjian W, Ying L. The influence of internet behavior on doctor-patient trust: based on CFPS 2016. J Northwest Norm Univ (Social Sci) 2019;56(2):119-126. [doi: 10.16783/j.cnki.nwnus.2019.02.016]
67. Bowen Z, Jiaojiang L. Will internet use affect public trust in doctors? An empirical analysis based on data CSS2013. Jiangsu Social Sci 2017(3):70-78. [doi: 10.13858/j.cnki.cn32-1312/c.2017.03.011]
68. Rui L. Research on the Influence of Online Medical Community Heaith Information Seeking Behavior on Patient Compliance. Beijing: Beijing Jiaotong University; 2019. URL: https://tinyurl.com/yckt32jk [accessed 2021-10-15]
69. Yiqun B. Influence of internet use on doctor-patient trust in the view of patients. Chin J Med Manag Sci 2021;11(04):78-83. [doi: 10.3969/j.issn.2095-7432.2021.04.017]
70. Hongze Z, Jingjing X, Suxiong W. Comparative analysis of the influence of media use on physician-patient trust and social trust. Journalism Mass Commun Mon 2017(06):68-76. [doi: 10.15897/j.cnki.cn51-1046/g2.2017.06.011]
71. Audrain-Pontevia A, Menvielle L. Effects of interpersonal trust among users of online health communities on patient trust in and satisfaction with their physician. Int J Technol Assess Health Care 2018 Feb 12;34(1):56-62. [doi: 10.1017/s0266462317004433]
72. Flynn KE, Smith MA, Freese J. When do older adults turn to the internet for health information? Findings from the Wisconsin Longitudinal Study. J Gen Intern Med 2006 Dec;21(12):1295-1301 [FREE Full text] [doi: 10.1111/j.1525-1497.2006.00622.x] [Medline: 16995892]
73. Chaudhuri S, Le T, White C, Thompson H, Demiris G. Examining health information-seeking behaviors of older adults. Comput Inform Nurs 2013 Nov;31(11):547-553 [FREE Full text] [doi: 10.1097/01.NCN.0000432131.92020.42] [Medline: 23974574]
74. Bylund CL, Gueguen JA, Sabee CM, Imes RS, Li Y, Sanford AA. Provider-patient dialogue about Internet health information: an exploration of strategies to improve the provider-patient relationship. Patient Educ Couns 2007 Jun;66(3):346-352. [doi: 10.1016/j.pec.2007.01.009] [Medline: 17329062]
75. Duckett J, Hunt K, Munro N, Sutton M. Does distrust in providers affect health-care utilization in China? Health Policy Plan 2016 Oct;31(8):1001-1009 [FREE Full text] [doi: 10.1093/heapol/czw024] [Medline: 27117483]
76. De Santis M, De Luca C, Quattrocchi T, Visconti D, Cesari E, Mappa I, et al. Use of the internet by women seeking information about potentially teratogenic agents. Eur J Obstet Gynecol Reprod Biol 2010 Aug;151(2):154-157. [doi: 10.1016/j.ejogrb.2010.04.018] [Medline: 20478650]
77. McMullan M. Patients using the internet to obtain health information: how this affects the patient-health professional relationship. Patient Educ Couns 2006 Oct;63(1-2):24-28. [doi: 10.1016/j.pec.2005.10.006] [Medline: 16406474]
78. Abdulla S, Vielhaber S, Machts J, Heinze H, Dengler R, Petri S. Information needs and information-seeking preferences of ALS patients and their carers. Amyotroph Lateral Scler Frontotemporal Degener 2014 Dec;15(7-8):505-512. [doi: 10.3109/21678421.2014.932385] [Medline: 25007829]
79. Schoenthaler A, Kalet A, Nicholson J, Lipkin M. Does improving patient-practitioner communication improve clinical outcomes in patients with cardiovascular diseases? A systematic review of the evidence. Patient Educ Couns 2014 Jul;96(1):3-12 [FREE Full text] [doi: 10.1016/j.pec.2014.04.006] [Medline: 24795073]
80. Jacob J. Consumer access to health care information: its effect on the physician-patient relationship. Alaska Med 2002;44(4):75-82. [Medline: 12650084]
81. Holland ML, Fagnano M. Appropriate antibiotic use for acute otitis media: what consumers find using Web searches. Clin Pediatr (Phila) 2008 Jun;47(5):452-456. [doi: 10.1177/0009922807313274] [Medline: 18310526]
82. Qi S, Huiling R, Yifan C, Wanjun X, Jinyin L. Survey on the status of health information use on the internet of community residents in Hefei. Med Soc 2014;27(06):62-64. [doi: 10.13723/j.yxysh.2014.06.020]
83. Neelapala P, Duvvi SK, Kumar G, Kumar BN. Do gynaecology outpatients use the internet to seek health information? A questionnaire survey. J Eval Clin Pract 2008 Apr;14(2):300-304. [doi: 10.1111/j.1365-2753.2007.00854.x] [Medline: 18324935]

Abbreviations

CNKI: China National Knowledge Infrastructure
MESH: Medical Subject Headings
OHC: online health community
OHI: online health information
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by G Eysenbach; submitted 10.08.20; peer-reviewed by K Lee, L Sbaffi, N Hirvonen; comments to author 21.10.20; revised version received 24.10.21; accepted 17.11.21; published 10.02.22.

Please cite as:
Luo A, Qin L, Yuan Y, Yang Z, Liu F, Huang P, Xie W
The Effect of Online Health Information Seeking on Physician-Patient Relationships: Systematic Review
J Med Internet Res 2022;24(2):e23354
URL: https://www.jmir.org/2022/2/e23354
doi:10.2196/23354
PMID:35142620

Review

The Effect of Digital Health Interventions on Parents' Mental Health Literacy and Help Seeking for Their Child's Mental Health Problem: Systematic Review

Daniel Peyton[1,2], MBBS, MPH, FRACP; Marquelle Goods[1], BA, Grad Dip; Harriet Hiscock[1,2,3], MBBS, FRACP, MD

[1]Murdoch Children's Research Institute, Parkville, Australia

[2]Department of Paediatrics, University of Melbourne, Melbourne, Australia

[3]Health Services Research Unit, The Royal Children's Hospital, Melbourne, Australia

Corresponding Author:
Harriet Hiscock, MBBS, FRACP, MD
Murdoch Children's Research Institute
50 Flemington Road
Parkville, 3052
Australia
Phone: 61 93456910
Email: harriet.hiscock@rch.org.au

Abstract

Background: Many children with mental health problems do not receive professional help. Despite the frequent use of digital health interventions (DHIs) such as websites or web-based service navigation platforms, their effects on parents' mental health literacy, help seeking, or uptake of professional services are unclear.

Objective: This study aims to provide a systematic review and narrative synthesis to describe whether DHIs improve the aforementioned parental outcomes.

Methods: Databases, including CINAHL, Embase, MEDLINE OVID, PsycINFO, and PubMed (2000-2020), were accessed. Studies were included if they evaluated quantitative changes in mental health literacy, help seeking, or the uptake of services by parents of children with mental health problems. Theoretical frameworks, sample sizes, participant demographics, recruitment, interventions, DHI use, results, and health economic measures were used for data extraction.

Results: Of the 11,379 search results, 5 (0.04%) studies met the inclusion criteria. One randomized controlled trial found the reduced uptake of services after using a DHI coupled with a telephone coach for a child's behavioral problem. Of 3 studies, 2 (66.7%) found statistically significant improvement in mental health literacy for attention-deficit/hyperactivity disorder but had no control group. One study found nonsignificant improvement in mental health literacy and help-seeking attitudes toward anxiety and depression compared with those in active controls. All studies were rated as having a high or serious risk of bias. Search results were affected because of a single reviewer screening articles, overall low-quality studies, and a lack of consistent nomenclature.

Conclusions: There is no high-quality evidence that DHIs can improve parents' mental health literacy, help seeking, or uptake of services. More research is needed to evaluate DHIs by using rigorous study designs and consistent measures.

Trial Registration: PROSPERO International Prospective Register of Systematic Reviews CRD42020130074; https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020130074

(J Med Internet Res 2022;24(2):e28771) doi:10.2196/28771

KEYWORDS

child; mental health; systematic review; caregiver; health literacy; digital health

Introduction

Background

Mental health problems are common among children [1,2]. They include internalizing problems, such as anxiety and lowered mood, and externalizing problems, such as hyperactivity, oppositional defiance, and aggression. Around half of these problems can progress to mental health disorders that are associated with adverse outcomes, including early school dropout, criminal justice system involvement, lower life satisfaction, poorer relationships, and lower earning potential [3-9]. Fortunately, there is a range of evidence-based treatments that have been shown to improve mental health problems in children, including the use of websites or web-based programs or other digital health interventions (DHIs) [10-13]. A DHI can be defined as the digital delivery of health information, such as through websites or apps, for health-related purposes [14]. Many of these treatments, including those delivered by DHIs and face-to-face interventions, focus on improving parenting—a key modifiable risk factor for these problems [15]. Despite treatments being available, many children with mental health problems do not receive professional help [2,16-18].

There are several recurrent barriers that prevent children receiving professional help. These barriers can be viewed along the help-seeking process, as parents need to recognize their child's problem and acknowledge their need for additional support, be aware of treatment options, overcome stigma in accessing treatment, and ultimately access available services or treatments [2,19,20]. A lack of problem recognition and awareness of available treatments reflect inadequate mental health literacy, which has been defined as the "knowledge and beliefs about mental disorders which aid their recognition, management or prevention" [21]. Mental health literacy is important because it is linked to actions and mental health outcomes [22]. For children, especially young children, parents play a large role in recognizing the child's problem and facilitating help seeking (Figure 1).

Ideally, we should be able to improve parents' knowledge of mental health problems in children and where to find available and accessible services to help their children. This could be done by improving their mental health literacy, a known modifiable factor of help seeking [23]. However, previous research on interventions designed to improve mental health literacy and help seeking has been hampered by a lack of consistent measures of mental health literacy and a lack of focus on parents [22,24,25]. For parents, a US study with 165 children with mood disorders and other mental health comorbidities showed that face-to-face mental health literacy interventions can improve the quality of services accessed by families compared with waitlist control. The quality of services was measured by consensus among a group of blinded expert clinician researchers [26]. However, this intervention was intensive (8 group sessions lasting 90 minutes each) and may have been affected by attrition bias, as only 74% of participants completed the 18-month follow-up. In addition, several families dropped out of the waitlist control group after their child's symptoms improved, underscoring the need for controlled trials to account for the natural history of some mental health problems improving over time.

Digital delivery of this educational material to parents, such as through a DHI, may prove to be an effective, accessible, scalable, and desirable way to improve parents' mental health literacy and help seeking. Most parents search the web for health information and seek out the lived experience of other parents through forums, such as those on Facebook [27,28]. As parents seek out this information on the web, money and resources are devoted to building websites, apps, and platforms to help parents better understand their child's mental health and where to receive help. Child mental health websites, such as childmind.org, can have enormous reach with a recent mental health campaign reaching 275 million people [29].

The World Health Organization states that DHIs have many perceived benefits, including enhanced reach, accessibility, scalability, desirability, reduced stigma, and perceived cost-effectiveness [14]. DHIs' perception of cost-effectiveness comes from the potential for near-infinite scalability at low cost and targeted early intervention [14,30,31]. However, data on cost-effectiveness are rarely collected, despite recommendations to measure the economic impact as part of any DHI evaluation [32,33].

DHIs have been shown to improve mental health literacy in adults, based on the findings of 2 systematic reviews [34,35]. However, these reviews, which included a combined total of 28 studies, only included 1 study with parents.

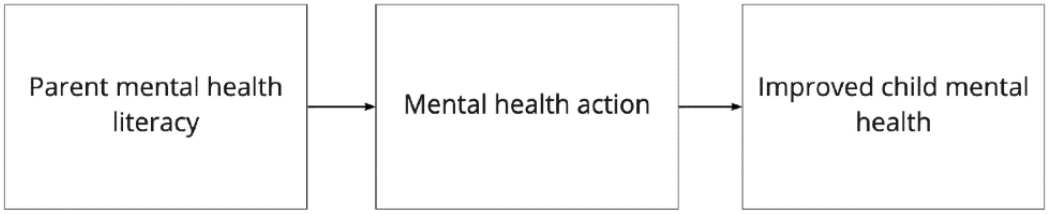
The single-parent study was a randomized controlled trial that found that a convenience sample of parents recruited from a single workplace improved their mental health literacy from a DHI [36]. This lack of focus on parents in previous reviews is important because parents are the agents of change for their child's mental health. Unlike adults seeking help for themselves, parents' willingness to receive help for their child's mental health problem is influenced by unique factors, such as whether the child participates in mental health treatment, or whether the treatment is framed in terms of child development [37,38]. With half of all adult mental health disorders originating in childhood, it is crucial to determine how DHIs can improve parents' mental health literacy, help seeking, and uptake of mental health services for their children [9].

However, there have been no consistent positive effects on parental help-seeking attitudes, with some low-quality studies finding a positive effect of DHIs, but most found no effects [34,35]. Studies in these 2 reviews had some limitations, specifically the common use of convenience sampling, the predominant focus on young people, lack of consistent measures, and low-quality evidence.

Recently, a universal education program delivered via SMS text messaging improved mental health literacy in the parents of adolescents compared with care as usual control. However, this study did not include parents of younger children or parents who were identified as having an adolescent with a mental health problem, who may be more likely to benefit from an intervention that facilitates help seeking [39].

Little is known about the effects of a DHI on the mental health literacy of parents, especially parents of young children, and even less is known about the effects on help seeking and uptake of services and cost-effectiveness. This is despite the frequent use of DHIs by parents and low uptake of services among many children with a mental health problem.

Figure 1. Link between parent mental health literacy and child mental health outcomes (adapted from a study by Jorm [22]).



Objectives

In this study, we aim to conduct a systematic review of the literature to understand (1) whether DHIs targeting parents of children aged 2 to 12 years with a mental health problem improve mental health literacy and (2) whether the use of DHIs is associated with changes in parental help seeking or uptake of mental health services for their child. We also aim to report the cost-effectiveness of such DHIs.

Methods

The systematic review was registered with PROSPERO (CRD42020130074). We conducted and reported a systematic review according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [40].

Eligibility

We included studies that evaluated a DHI delivered directly to parents of children aged 2 to 12 years, with quantitative data reporting on outcomes of mental health literacy (specifically knowledge of treatment), help seeking (attitudes, intentions, and behaviors), or uptake of mental health services. Quantitative data were chosen to narratively synthesize the impact of DHIs on mental health literacy, help seeking and uptake of services.

For this review, we defined a DHI as a consumer-facing intervention using information communication technology targeting parents. The intervention could deliver information as a static webpage, a web-based parenting program, a web-based social network, a native mobile app, or other content delivered using digital means (other than telehealth). This definition was included in the PROSPERO registration.

We included DHIs targeting children with and without a mental health condition as long as the DHI was delivered as part of a program where some families were identified as having a mental health concern for their child. We included children aged 2 to 12 years. This age range was selected because of their likely dependence on parents to receive help for their mental health and the long-term impact of these early years on the well-being into adulthood [9]. We required a minimum of 1 outcome question on mental health literacy focusing on any of the following: knowledge of treatment, help seeking, or uptake of services.

Study designs included randomized controlled trials, quasi-randomized trials, and uncontrolled single-cohort studies. We restricted our analysis to articles published between January 2000 and December 2020 and written in English. We excluded conference proceedings and gray literature.

Data Sources and Search Strategy

We developed our search strategy after consultation with a research librarian at The Royal Children's Hospital, Melbourne, Australia. A pilot search was performed in MEDLINE OVID, followed by a review of keywords and further development of the search strategy. We searched the electronic databases CINAHL, Embase, MEDLINE OVID, PsycINFO, and PubMed in late 2019 and repeated the search in January 2021 to identify any more recent publications.

We also reviewed the reference list of the included studies to identify additional studies for full-text review. All search results were compiled in Endnote and then exported to Covidence for screening. The search strategy used for all the databases is given in Multimedia Appendix 1.

Study Selection

One author (DP) screened the titles and abstracts of all articles produced from the search against the eligibility criteria. The full text of the remaining articles was obtained and screened again against the inclusion criteria. Any concerns about study eligibility were resolved in discussions with the supervising author (HH) during fortnightly supervision meetings. If there was insufficient evidence from the full-text study on whether it met the inclusion or exclusion criteria, DP attempted to contact the authors to obtain relevant information.

Data Collection Process

Two authors (DP and MG) independently extracted data from the included studies using a pre-existing data collection form for intervention reviews from Cochrane [41].

Data Items

Data extracted included study design; number of participants; type of comparison (where relevant); setting; recruitment; age and sex of participants and their children; the intervention, including the theoretical basis (a factor that may influence the success of a help-seeking intervention) [42] and measures of DHI use; outcome measures and whether they are validated measures; results; and economic outcomes. The data extracted were compared for accuracy, and the supervising author (HH) resolved any disagreements. Where possible, we calculated the effect sizes of the interventions and included these in Table 1.

Table 1. Primary outcomes of the interventions.

Study	Design	Sample, n	Intervention	Timing of measures	Primary outcome	Measure	Outcome	*P* value	Validated measure
Montoya et al [43]	Pre or post single cohort	35	DISCERN tool assessing popular Spanish websites about ADHD[a] treatment	Unspecified time points pre, post parents using the DISCERN tool	Mental health literacy: ADHD specific knowledge and motivation for treatment	• The ADHD-knowledge and motivation for treatment questionnaire (ADHD-KMT). • Basic knowledge subscale	• Pre: mean 49.09 (SD 9.46) • Post: mean 63.21 (SD 9.45) • Cohen *d*=1.49	• <.01	No
Ossebaard et al [44]	Pre or post single cohort	195	Web-based decision aid on ADHD treatment	Pre, post intervention, though exact timing unclear	Mental health literacy: ADHD knowledge and treatment	• "Would you please rate your knowledge on ADHD and its treatment possibilities" with a response on a 1-10 numerical scale	• Pre: mean 6.2 (SD 1.9) • Post: mean 6.5 (SD 1.9) • Cohen *d*=0.16	• .60	Unclear
Ryan et al [45]	Pre or post single cohort	172	Information-based website on ADHD management	Baseline: 28 days post-baseline	Mental health literacy: ADHD knowledge	• ADHD Knowledge and Opinions Survey-Revised (AKOS-R) – adapted • Lower score (min: 30; max: 60)=higher knowledge	• Wilcoxon signed rank test showed a statistically significant moderate increase in knowledge; Z=−4.799; Cohen *d*=−0.503	• <.01	No

Study	Design	Sample, n	Intervention	Timing of measures	Primary outcome	Measure	Outcome	*P* value	Validated measure
Sapru et al [46]	Nonrandomized controlled trials	27	3× PowerPoint presentations emailed to participants	Pre and post intervention, though exact timing unclear	Mental health literacy and help-seeking attitudes for depression	• Understanding mood disorders questionnaire • Lower incorrect score=higher knowledge	• Median number of incorrect scores: Intervention: Pre 7, post 1; Control: Pre 7.5, post 4 • Within-group difference (pre or post) in PowerPoint group: Wilcoxon signed-rank test showed statistically significant improvement in responses (Z=−2.30; *P*=.04) • Comparison between PowerPoint group and control (in-person group): One-way ANOVA showed no statistically significant improvement difference in responses	• Within-group difference (pre or post) in PowerPoint group: *P*=.04 • Comparison between PowerPoint group and control (in-person group): *P* value not reported	Not reported
				Pre, post intervention, though exact timing unclear	Mental health literacy and help-seeking attitudes for anxiety	• Understanding of anxiety disorders questionnaire • Lower incorrect score=higher knowledge		• Within-group difference (pre or post) in PowerPoint group: *P*=.04 • Comparison between Power Point group and control (in-person group): *P* value not reported	Not reported

Study	Design	Sample, n	Intervention	Timing of measures	Primary outcome	Measure	Outcome	*P* value	Validated measure
							• Median number of incorrect scores: Intervention: Pre 9, post 2; Control: Pre 6.5, post 3.5 • Within-group difference (pre or post) in PowerPoint group: Wilcoxon signed-rank test showed statistically significant improvement in responses (Z=−2.30, *P*=.04) • Comparison between PowerPoint group and control (in-person group): one-way ANOVA showed no statistically significant improvement difference in responses		
Sourander et al [47]	Randomized controlled trial	464	Strongest Families' Smart website and 11× weekly 45-minute telephone coaching sessions	6 months, 12 months, and 2 years after randomization	Uptake of services in the past 6 months	• Past service use evaluated using a yes or no question: "asking the parents if the child had received any behavioural treatment in the last 6 months"	• Number of participants reporting uptake of services: Intervention: 28 (18%); Control: 46 (28%); OR 1.8 [95% CI 1.1-3.1]	• .02	No

[a]ADHD: attention-deficit/hyperactivity disorder.

Risk of Bias

The included studies were assessed for quality against 1 of 2 instruments. For nonrandomized studies, we assessed the risk of bias using the Risk of Bias in Nonrandomized Studies of Interventions tool [48]. For randomized studies, we assessed bias using the revised Cochrane tool for assessing risk of bias in randomized trials [49]. The quality assessment was conducted independently by DP and MG. They compared their assessments and resolved any disputes by discussion or through the input of the supervising author (HH).

Summary Measures

Whenever possible, we presented the outcome data of mental health literacy, help seeking, and uptake of services consistently, with parametric continuous data compared using means, nonparametric continuous data presented using medians, and categorical data presented as proportions. We also attempted to group the outcome data by validated and unvalidated measures.

Synthesis

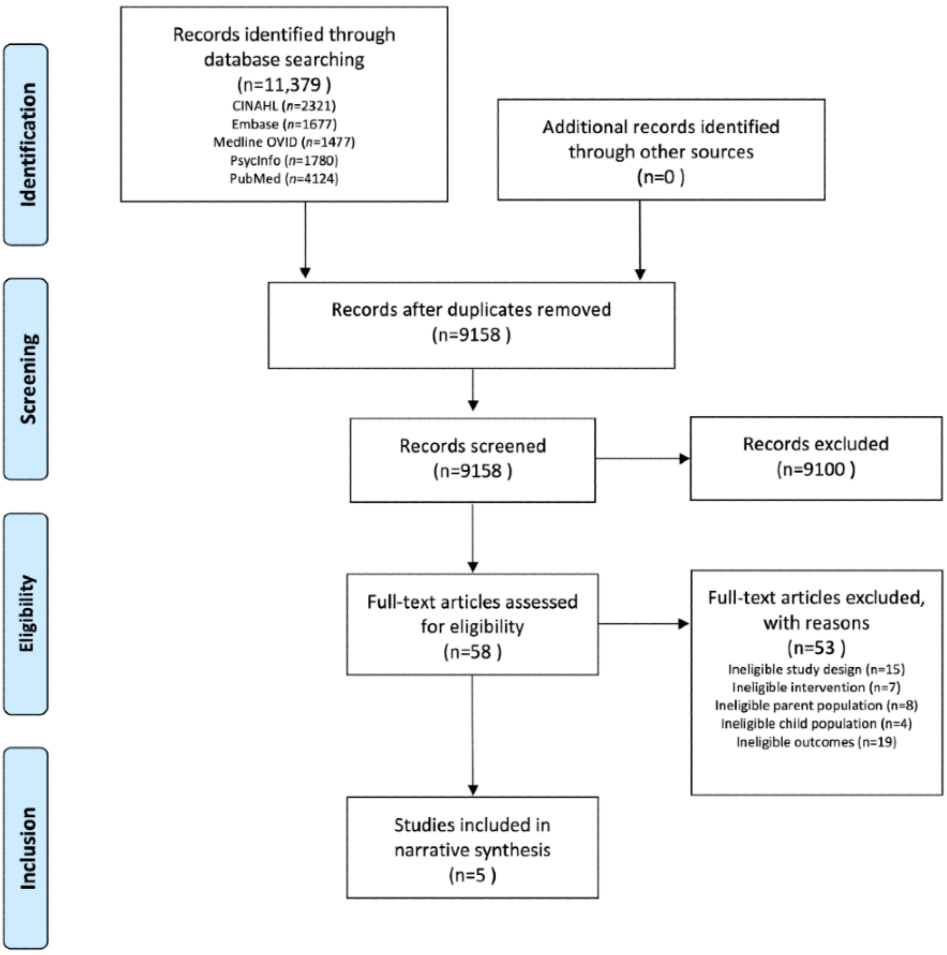
Owing to the heterogeneity in outcome measures, we could not conduct a meta-analysis. Accordingly, we used a narrative synthesis to describe the effects of the DHIs.

Results

Search Results

Through the search strategy detailed in the previous section, a total of 11,379 potentially eligible articles were identified. Of the 11,379 articles, 5 (0.04%) met all inclusion and exclusion criteria (Figure 2). The primary author (DP) reviewed the reference list of these included studies, which revealed no additional studies meeting inclusion and exclusion criteria.

Figure 2. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of search results and study selection.



Description of Included Studies

Of the 5 included studies, 1 (20%) was a randomized controlled trial [47], 1 (20%) was a nonrandomized trial of 2 interventions [46], and 3 (60%) were uncontrolled before and after studies [43-45] (Table 2).

All 5 studies were published between 2010 and 2018. There were 893 participants across the 5 studies, with the number of participants ranged from 27 to 464. The mean age of the children ranged from 4 to 10 years across the 5 studies. All studies were published in Europe or North America. Outcome measures included knowledge of attention-deficit/hyperactivity disorder (ADHD) treatment, knowledge and help seeking for anxiety and depression, and uptake of treatment for a child's behavioral problem.

A total of 3 studies included participants with concerns about, or a recent diagnosis of, ADHD [43-45]. Another study included only participants who were parents of children with high-level disruptive behavior on the Strengths and Difficulties Questionnaire and who recognized that their child had a problem [47]. The final study included participants who were referred to a tertiary center for management of the child's anxiety or depressive disorder, although the authors did not describe how the disorder had been diagnosed [46].

Participants were sampled using a variety of techniques. Of the 5 studies, 2 (40%) used consecutive sampling techniques to approach participants attending a scheduled universal health appointment [47] or a tertiary hospital mental health outpatient clinic [46]; 1 (20%) used a convenience sample of participants who had already attempted to access the intervention evaluated in the study [44]; and 1 (20%) used a convenience sample in which participants were selected by their child's physician or from a local advocacy group [43].

Table 2. Study description.

Study	Country	Design	Sample, n	Participants	Recruitment	Intervention	Comparator	Theoretical basis for the intervention	Digital health intervention use	Economic outcomes
Montoya et al [43]	Spain	Single cohort pre or post study	35	Parents of children with a recent diagnosis of ADHD[a]	Parents selected by their child's physician or from a local advocacy group	Use of the DISCERN tool to assess the quality of 10 popular Spanish websites about ADHD treatment	Nil	Not reported	• Not reported	Not reported
Ossebaard et al [44]	Netherlands	Single cohort pre or post study	195	Parents of children with a recent diagnosis of ADHD	The web-based decision aid invited visitors to the website to participate in the study	Web-based decision aid on ADHD treatment	Nil	Yes	• About 7500 unique visits • About 6 minutes on site • About 8-9 clicks to navigate	Not reported
Ryan et al [45]	United Kingdom	Single cohort pre or post study	172	Parent or carer of a child with confirmed or suspected ADHD	Invited to attend if attending one of 3 pediatric outpatient clinics for suspected or confirmed ADHD	Information based website on ADHD management	Nil	Not reported	• Never used the website: 62 (41%) • 1-2 times: 50 (33%) • 4-5 times: 27 (18%) • 5-6 times: 6 (4%) • 7+ times: 8 (5%)	Not reported
Sapru et al [46]	Canada	Prospective nonrandomized controlled trial before and after study	27	Families referred to a tertiary hospital for management of a mood or anxiety disorder	Families on a waitlist for outpatient treatment of depression or anxiety were invited to attend	3× PowerPoint presentations emailed to participants	3 × 1-hour in-person group family psychoeducation sessions	Yes	• PowerPoint presentations completed: mean 2.7 (SD 2.7) • Control group: mean 3.75 (SD 2.3)	Not reported
Sourander et al [47]	Finland	Prospective randomized controlled trial	464	Parents of children with high level disruptive behavior at a universal 4-year-old health check	Families attending a universal 4-year-old health check were screened and invited to attend	Strongest families smart website and 11× weekly 45-minute telephone coaching sessions	Brief website on positive parenting strategies and single 45-minute telephone coaching session and standard care	Not reported	• Not reported	Not reported

[a]ADHD: attention-deficit/hyperactivity disorder.

Description of the Included Interventions

Of the 5 interventions, 4 (80%) were delivered on the web through a website [43-45,47] and 1 (20%) was delivered via a series of PowerPoint presentations [46]. These PowerPoint presentations were emailed to each family every week for 3 weeks. The topics of the three PowerPoint presentations were (1) introduction and treatment options, (2) interpersonal illness and communication skills, and (3) problem solving and personal reflection [46].

A total of 2 (40%) web-based interventions were delivered with a cointervention [43,47]. One (20%) of these cointerventions consisted of 11 consecutive weekly telephone coaching sessions, in addition to access to the Strongest Families Smart Website [47]. This website features 11 sessions containing tailored content, exercises, and instructional videos and requires parents to complete knowledge and experience-based questions. This content is designed to help parents develop skills to promote positive behavior and a positive relationship with their children [47]. Another study by Montoya et al [43] used a cointervention. In this study, parents evaluated popular ADHD websites against the DISCERN instrument [50] to assess the quality of written consumer health information available on ADHD treatment [43].

The remaining 2 (40%) interventions consisted of a website focused on ADHD [44,45]. A study by Ossebaard et al [44] trialed a web-based decision aid designed to help support parents and caregivers through the decision-making process of ADHD treatment. The average visitor, which included participants and nonparticipants, visited the website for an average of 6 minutes [44]. The final ADHD website contained information on the management of ADHD [45]. The website was funded by the pharmaceutical company Shire, which was disclosed to the participants. The participants could access the website for 1 month, and most of the participants accessed the website once or twice during that time [45]. For these 2 ADHD websites, postintervention outcomes were measured immediately following the intervention [44], 30 days after the intervention started [45], or 2 years after the intervention commenced [47].

Of the 5 studies, 2 (40%) did not specify precisely when they recorded postintervention outcomes [43,46].

Effect on Mental Health Literacy, Help Seeking, and Uptake of Services

Mental health literacy outcomes were the most common outcome assessed by the included studies, with 80% (4/5) of the studies measuring some form of mental health treatment knowledge (Table 1). The most common mental health problem assessed by the knowledge measures was ADHD [43-45], followed by depression and anxiety knowledge and help-seeking attitudes studied by Sapru et al [46]. Only 1 (20%) study measured the parent-reported uptake of mental health services [47].

ADHD Knowledge

Despite 60% (3/5) of the studies intending to measure ADHD knowledge and all through survey responses, each study used a different measure. None of these measures were validated. An adapted version of a validated measure was used by Ryan et al [45], but the authors did not provide a description of how it had been adapted and whether it was still valid. All of the ADHD knowledge studies were uncontrolled pre-post studies, and all showed an improvement in parent ADHD knowledge scores, 2 (40%) of which were statistically significant [42,49].

In addition, changes in knowledge among those who accessed the website and those who did not were assessed by Ryan et al [45]. Their study [45] showed that those who accessed the website at least once had a moderately significant improvement in knowledge compared with those who never accessed the *ADHD and You* website.

Of note, evaluation of a web-based decision aid by Ossebaard et al [44] was affected by a large number of missing data. From the 7500 unique views to the site, all of whom were invited to participate in the study, only 195 participants were enrolled, leading to potential selection bias. In addition, of these 195 participants, only 12 (6.2%) provided outcome data before and after the intervention, leading to potential attrition bias.

Depression and Anxiety Knowledge and Attitudes to Help Seeking

The only study that evaluated anxiety and depression-based mental health literacy and help-seeking attitudes was carried out by Sapru et al [46]. One measure was used for anxiety, and another for depression, with each measure assessing both knowledge and help-seeking attitudes within the same instrument. It was not reported whether these tools had been validated for this population.

Both the anxiety and depression measures showed an improvement in median scores of the intervention (web-based) compared with those of the control (in-person) group, although this difference was not significant in a small sample size.

Missing data and high attrition rates were again common, with outcome data provided for only 38% (5/13) of the intervention participants and 57% (8/14) of the control participants. The authors did not report why so many families failed to initiate or complete the programs and outcome measures. Two of the authors were contacted but did not provide further clarification on reasons for the missing data.

Uptake of Mental Health Services

Only 1 (20%) study measured the uptake of mental health services, which was also the largest study and had the longest follow-up of 2 years [47]. A study by Sourander et al [47] asked parents to self-report whether they had received any behavior treatment for their child in the previous 6 months. This measure was recorded at 6, 12, and 24 months after starting the 11-week intervention. The authors did not report whether this measure had been validated. Fewer parents in the intervention group, consisting of a website and 11 weekly telephone coaching sessions, reported that their child had accessed behavioral treatments (28/160, 17.5% participants) than did parents in the control group (46/164, 28% participants; odds ratio 1.8, 95% CI 1.1-3.1; P=.02). This reduction in the uptake of behavioral treatments occurred in the context of a small but significant

improvement in the child's behavior in the intervention group compared with the control group.

Cost-effectiveness

No studies reported on the cost-effectiveness or costs of the DHIs.

Assessment of Risk of Bias

One randomized controlled trial was rated as having a high risk of bias in 1 domain because of missing data, giving it an overall rating of high risk (Table 3) [47].

A total of 4 study designs were nonrandomized, with 3 (75%) of these studies [43,45,46] rated at serious risk of bias and 1 (25%) [44] rated at critical risk of bias (Table 4). The studies were rated at serious risk of bias because of a lack of identification of, or control for, potential confounders; potential for bias in selection of participants; and lack of objective outcome measures. The large number of missing participants also contributed to attrition bias and subsequent critical risk ratings.

Table 3. Risk of bias of randomized studies using the Cochrane tool for assessing risk of bias in randomized trials (RoB 2).

Study	Randomization process or selection bias	Deviations from intervention	Missing outcome or attrition bias	Measurement of outcome or detection bias	Selection of reported result or reporting bias	Overall
Sourander et al [47]	Low	Low	High	Some concerns	Some concerns	High

Table 4. Risk of bias in nonrandomized studies using the Risk of Bias in Non-randomized Studies–of Interventions (ROBINS-I) tool.

Study	Confounding	Selection of participants	Classification of interventions	Deviations from intended interventions	Missing data	Measurement of outcomes	Selection of reported result	Overall
Montoya et al [43]	Serious	Low	Low	Low	Low	Moderate	Moderate	Serious
Ossebaard et al [44]	Serious	Critical	Low	Low	Critical	Serious	Moderate	Critical
Ryan et al [45]	Serious	Low	Low	Low	Moderate	Serious	Moderate	Serious
Sapru et al [46]	Serious	Serious	Low	Low	Moderate	Moderate	Moderate	Serious

Discussion

Principal Findings

This study identified 5 studies of DHIs for parents of children with a mental health problem, measuring changes in mental health literacy, help seeking, or uptake of services.

Of those measuring mental health literacy, 80% (4/5) of the studies showed an improvement in parent knowledge. However, most of these studies focused on ADHD knowledge and were of low quality.

Of the 5 studies, 1 (20%), using a very small sample size of parents, measured both mental health literacy and help-seeking attitudes and used a nonrandomized control group, showing a nonsignificant trend to improved knowledge and help-seeking attitudes for child's anxiety and depression. For this study, the mental health literacy and help-seeking attitudes outcomes were evaluated using the same measure and results were not presented separately, precluding conclusions about whether this improvement was predominantly because of changes in knowledge or attitudes.

The only large randomized controlled trial measured uptake of services and found the use of a website coupled with a telephone coach, reduced uptake of services for the child's behavior, whilst simultaneously improving child behavior compared with a control group at 24 months follow up [47]. Despite the widespread use of websites and apps to help parents understand their child's mental health or find services to help their child, only one study evaluated a universally accessible website [43]. Of the 5 studies, 2 (40%) had a comparison group, and neither of these studies compared the DHI to an existing and previously evaluated face-to-face, web-based, or school-based intervention. Thus, the comparative efficacy, feasibility, and cost-effectiveness of DHIs and face-to-face interventions remain unclear.

Of the 5 studies, 2 (40%) reported using theory to inform the design of the DHI. Although there is no evidence to definitively support the use of theory in designing a DHI, it is recommended to use a theory, or theories, to inform the design of health promotion interventions, and it may be beneficial for DHIs targeting help seeking [42,51].

None of the studies reported health economic outcomes of the interventions, such as development costs, implementation expenses, or potential financial benefits from the intervention on the family or health services. The overall quality of the papers was low, with only 20% (1/5) of the studies being a randomized controlled trial. All studies were rated as either high risk of bias on the revised Cochrane tool or serious or critical risk of bias on the Risk of Bias in Nonrandomized Studies of Interventions tool.

In addition, the lack of consistent and validated measures made a meta-analysis impossible and limited our ability to compare efficacy among the interventions. The lack of consistent measures has been described previously [24].

This is the only review showing the impact of DHIs on mental health literacy, help seeking, and referral uptake in parents of children with mental health problems. We searched a wide range of databases, hand searched references from included articles, and attempted to contact authors where data were missing. This study included all quantitative studies evaluating a DHI across multiple time points and thus presented a wider scope of included study designs than existing review articles on DHIs for mental health literacy or help seeking. Finally, this was the only study that extracted data on the theoretical basis of the intervention and economic outcomes.

Limitations

We included only studies with quantitative outcome measures. We recognize that we could have used categorical coding of qualitative data (eg, positive, neutral, or negative impact) to include qualitative research. This could be an area for future research. In addition, qualitative studies may provide more nuanced data into the effectiveness or otherwise of DHIs in this area, particularly on factors influencing help-seeking attitudes, intentions, and behaviors. In addition, a single reviewer (DP) evaluated all search results against the inclusion and exclusion criteria, which may have resulted in studies being missed at the screening stage. However, hand searching of references within these papers revealed no new studies, suggesting that it is unlikely that we missed any published studies. The studies included were of poor quality; therefore, the results must be interpreted with caution. This review only included peer-reviewed journals and did not include a search for gray literature. As such, there is potential for publication bias in the results. Finally, a lack of consistent nomenclature around help seeking and uptake of services may have resulted in the search strategy missing some studies that measured these outcomes.

Impact

There is no high-quality evidence that DHIs improve parent mental health literacy, help seeking, or uptake of services, even for the most studied area of ADHD. There is low-quality evidence that parents' mental health literacy can be improved through the use of DHIs. There is also evidence that the use of a website and telephone coach may reduce the long-term uptake of mental health services for preschool children with disruptive behavior. The economic benefit of any DHI targeting parent mental health literacy, help seeking, or uptake of services remains unknown. This study cautiously supports the use of DHIs, especially ADHD websites, to improve parent mental health literacy. There is no evidence that any DHI can improve help seeking or uptake of services for children with a mental health problem.

Future Research

Despite the widespread availability, enthusiasm for, and use of DHIs among parents, there is little rigorous evidence regarding the effect of DHIs on parent mental health literacy, help seeking, and uptake of services for their children. There is an urgent need to develop, implement, and rigorously evaluate DHIs designed to improve these outcomes, including an economic evaluation of their effects. Websites targeting parent mental health literacy, especially for mental health problems other than ADHD, should be evaluated to establish whether they increase mental health literacy. Ideally, this evaluation would compare new and previously evaluated interventions using validated measures of parent mental health literacy.

Researchers should conduct randomized controlled trials of new and existing DHIs, including existing interventions that are already frequently accessed by parents. Comparison of face-to-face and school- or community-based interventions would also prove helpful in understanding the role of DHIs within the broader context of child mental health services [25]. Outcomes should include validated measures of parents' knowledge of mental health problems in children and mental health actions, such as help seeking and uptake of services [25]. Consistent use of validated measures would allow a comparison of interventions and meta-analysis of their effects [52]. Research focusing on help seeking and uptake of services is especially important, given that so many children with mental health disorders are not receiving professional help. Until such research is conducted, we do not know whether a DHI can improve the uptake of mental health services among parents of children with mental health problems. A systematic review of qualitative studies may provide additional information on the influence of DHIs on parents' help-seeking behaviors.

Conclusions

This review found low-quality evidence that DHIs may increase mental health literacy for ADHD and increase mental health literacy and help-seeking attitudes toward anxiety and depression. Overall, the heterogeneity of measures and high risk of bias across studies impacted our ability to confidently interpret these findings. We highlight the gap between parents' frequent use of web-based sources of health information and the paucity of published evidence on the effect of these DHIs on help seeking, the uptake of services, and cost-effectiveness.

Acknowledgments

The authors would like to thank Ms Poh Chua, a research librarian at The Royal Children's Hospital, Melbourne, Australia, for her help in developing the search strategy for this review. In addition, they would like to thank Associate Professor Emma Sciberras for her assistance in conceptualizing this paper. DP was funded by the National Health and Medical Research Council Postgraduate Scholarship (grant 1189935). HH was supported by the National Health and Medical Research Council Practitioner Fellowship Award 1136222. The Victorian Government's Operational Infrastructure Support Program supported research at the Murdoch Children's Research Institute.

Authors' Contributions

DP conceptualized and designed the study, searched the literature, extracted data, conducted a quality appraisal, drafted the initial manuscript, and reviewed and revised the manuscript. MG extracted data, conducted a quality appraisal, and reviewed and revised the manuscript. HH conceptualized and designed the study, supervised the data extraction and quality appraisal, and reviewed and revised the manuscript. All authors approved the final manuscript as submitted and agreed to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File , 20 KB - jmir_v24i2e28771_app1.docx]

References

1. Perou R, Bitsko RH, Blumberg SJ, Pastor P, Ghandour RM, Gfroerer JC, Centers for Disease Control & Prevention (CDC). Mental health surveillance among children--United States, 2005-2011. MMWR Suppl 2013 May 17;62(2):1-35. [Medline: 23677130]
2. Lawrence D, Johnson S, Hafekost J. The Mental Health of Children and Adolescents. Report on the Second Australian Child and Adolescent Survey of Mental Health and Wellbeing. Canberra: Department of Health; 2015.
3. Asendorpf JB, Denissen JJ, van Aken MA. Inhibited and aggressive preschool children at 23 years of age: personality and social transitions into adulthood. Dev Psychol 2008 Jul;44(4):997-1011. [doi: 10.1037/0012-1649.44.4.997] [Medline: 18605830]
4. Duchesne S, Vitaro F, Larose S, Tremblay RE. Trajectories of anxiety during elementary-school years and the prediction of high school noncompletion. J Youth Adolescence 2007 Sep 19;37(9):1134-1146. [doi: 10.1007/s10964-007-9224-0]
5. Knapp M, McDaid D, Parsonage M. Mental Health Promotion and Mental Illness Prevention: the Economic Case. London, UK: Department of Health; 2011.
6. Patel V, Flisher AJ, Hetrick S, McGorry P. Mental health of young people: a global public-health challenge. The Lancet 2007 Apr;369(9569):1302-1313. [doi: 10.1016/S0140-6736(07)60368-7]
7. Powdthavee N, Vignoles A. Mental health of parents and life satisfaction of children: a within-family analysis of intergenerational transmission of well-being. Soc Indic Res 2007 Nov 30;88(3):397-422. [doi: 10.1007/s11205-007-9223-2]
8. Stevenson J, Goodman R. Association between behaviour at age 3 years and adult criminality. Br J Psychiatry 2001 Sep;179:197-202. [doi: 10.1192/bjp.179.3.197] [Medline: 11532795]
9. Kessler RC, Amminger GP, Aguilar-Gaxiola S, Alonso J, Lee S, Ustün TB. Age of onset of mental disorders: a review of recent literature. Curr Opin Psychiatry 2007 Jul;20(4):359-364 [FREE Full text] [doi: 10.1097/YCO.0b013e32816ebc8c] [Medline: 17551351]
10. Furlong M, McGilloway S, Bywater T, Hutchings J, Smith S, Donnelly M. Behavioural and cognitive-behavioural group-based parenting programmes for early-onset conduct problems in children aged 3 to 12 years. Cochrane Database Syst Rev 2012 Feb 15(2):CD008225. [doi: 10.1002/14651858.CD008225.pub2] [Medline: 22336837]
11. Tarver J, Daley D, Lockwood J, Sayal K. Are self-directed parenting interventions sufficient for externalising behaviour problems in childhood? A systematic review and meta-analysis. Eur Child Adolesc Psychiatry 2014 Dec;23(12):1123-1137. [doi: 10.1007/s00787-014-0556-5] [Medline: 24842197]
12. March S, Spence SH, Donovan CL. The efficacy of an internet-based cognitive-behavioral therapy intervention for child anxiety disorders. J Pediatr Psychol 2009 Jun;34(5):474-487. [doi: 10.1093/jpepsy/jsn099] [Medline: 18794187]
13. James AC, James G, Cowdrey FA, Soler A, Choke A. Cognitive behavioural therapy for anxiety disorders in children and adolescents. Cochrane Database Syst Rev 2015 Feb 18(2):CD004690 [FREE Full text] [doi: 10.1002/14651858.CD004690.pub4] [Medline: 25692403]
14. World Health Organization. WHO Guideline: Recommendations on Digital Interventions for Health System Strengthening. Geneva: World Health Organization; 2018.
15. Kiernan KE, Mensah FK. Poverty, family resources and children's early educational attainment: the mediating role of parenting. British Educ Res J 2011 Apr;37(2):317-336. [doi: 10.1080/01411921003596911]
16. Mental health of children and young people in England, 2017. NHS Digital. 2018. URL: https://digital.nhs.uk/data-and-information/publications/statistical/mental-health-of-children-and-young-people-in-england/2017/2017 [accessed 2022-02-01]
17. Kataoka SH, Zhang L, Wells KB. Unmet need for mental health care among U.S. children: variation by ethnicity and insurance status. Am J Psychiatry 2002 Sep;159(9):1548-1555. [doi: 10.1176/appi.ajp.159.9.1548] [Medline: 12202276]
18. Mental health: A Report of the Surgeon General. Rockville, MD: Department of Health and Human Services, U.S. Public Health Service; 1999.

19. Radez J, Reardon T, Creswell C, Lawrence PJ, Evdoka-Burton G, Waite P. Why do children and adolescents (not) seek and access professional help for their mental health problems? A systematic review of quantitative and qualitative studies. Eur Child Adolesc Psychiatry 2021 Feb;30(2):183-211 [FREE Full text] [doi: 10.1007/s00787-019-01469-4] [Medline: 31965309]
20. Rickwood D, Deane FP, Wilson CJ, Ciarrochi J. Young people's help-seeking for mental health problems. Australian e J Advanc Mental Health 2014 Dec 17;4(3):218-251. [doi: 10.5172/jamh.4.3.218]
21. Jorm AF, Korten AE, Jacomb PA, Christensen H, Rodgers B, Pollitt P. "Mental health literacy": a survey of the public's ability to recognise mental disorders and their beliefs about the effectiveness of treatment. Med J Aust 1997 Feb 17;166(4):182-186. [doi: 10.5694/j.1326-5377.1997.tb140071.x] [Medline: 9066546]
22. Jorm A. We need to move from 'mental health literacy' to 'mental health action'. Mental Health Prevent 2020 Jun;18:200179. [doi: 10.1016/j.mhp.2020.200179]
23. Xu Z, Huang F, Kösters M, Staiger T, Becker T, Thornicroft G, et al. Effectiveness of interventions to promote help-seeking for mental health problems: systematic review and meta-analysis. Psychol Med 2018 Dec;48(16):2658-2667. [doi: 10.1017/S0033291718001265] [Medline: 29852885]
24. Hurley D, Swann C, Allen MS, Ferguson HL, Vella SA. A systematic review of parent and caregiver mental health literacy. Community Ment Health J 2020 Jan;56(1):2-21. [doi: 10.1007/s10597-019-00454-0] [Medline: 31541315]
25. Tully LA, Hawes DJ, Doyle FL, Sawyer MG, Dadds MR. A national child mental health literacy initiative is needed to reduce childhood mental health disorders. Aust N Z J Psychiatry 2019 Apr;53(4):286-290. [doi: 10.1177/0004867418821440] [Medline: 30654614]
26. Mendenhall AN, Fristad MA, Early TJ. Factors influencing service utilization and mood symptom severity in children with mood disorders: effects of multifamily psychoeducation groups (MFPGs). J Consult Clin Psychol 2009 Jun;77(3):463-473. [doi: 10.1037/a0014527] [Medline: 19485588]
27. Park E, Kim H, Steinhoff A. Health-related internet use by informal caregivers of children and adolescents: an integrative literature review. J Med Internet Res 2016 Mar 03;18(3):e57 [FREE Full text] [doi: 10.2196/jmir.4124] [Medline: 26940750]
28. Yardi S, Caldwell PH, Barnes EH, Scott KM. Determining parents' patterns of behaviour when searching for online information on their child's health. J Paediatr Child Health 2018 Nov;54(11):1246-1254. [doi: 10.1111/jpc.14068] [Medline: 29864197]
29. 2020 Children's Mental Health Report. Child Mind Institute. URL: https://childmind.org/awareness-campaigns/childrens-mental-health-report/2020-childrens-mental-health-report/ [accessed 2022-02-01]
30. Christensen H, Reynolds J, Griffiths KM. The use of e-health applications for anxiety and depression in young people: challenges and solutions. Early Interv Psychiatry 2011 Feb;5 Suppl 1:58-62. [doi: 10.1111/j.1751-7893.2010.00242.x] [Medline: 21208393]
31. Boydell KM, Hodgins M, Pignatiello A, Teshima J, Edwards H, Willis D. Using technology to deliver mental health services to children and youth: a scoping review. J Can Acad Child Adolesc Psychiatry 2014 May;23(2):87-99 [FREE Full text] [Medline: 24872824]
32. Hilty DM, Yellowlees PM, Nesbitt TS. Evolution of telepsychiatry to rural sites: changes over time in types of referral and in primary care providers' knowledge, skills and satisfaction. Gen Hosp Psychiatry 2006;28(5):367-373. [doi: 10.1016/j.genhosppsych.2006.05.009] [Medline: 16950370]
33. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
34. Tay JL, Tay YF, Klainin-Yobas P. Effectiveness of information and communication technologies interventions to increase mental health literacy: a systematic review. Early Interv Psychiatry 2018 Dec 13;12(6):1024-1037. [doi: 10.1111/eip.12695] [Medline: 29897166]
35. Brijnath B, Protheroe J, Mahtani KR, Antoniades J. Do web-based mental health literacy interventions improve the mental health literacy of adult consumers? Results from a systematic review. J Med Internet Res 2016 Jun 20;18(6):e165 [FREE Full text] [doi: 10.2196/jmir.5463] [Medline: 27323907]
36. Deitz DK, Cook RF, Billings DW, Hendrickson A. A web-based mental health program: reaching parents at work. J Pediatr Psychol 2009 Jun;34(5):488-494 [FREE Full text] [doi: 10.1093/jpepsy/jsn108] [Medline: 18845585]
37. Perceptions of parenting: mapping the gaps between expert and public understandings of effective parenting in Australia. Frameworks Institute. URL: https://www.frameworksinstitute.org/wp-content/uploads/2020/06/PRC_MTG_Report_May_2016_final.pdf [accessed 2022-02-01]
38. Cunningham CE, Chen Y, Deal K, Rimas H, McGrath P, Reid G, et al. The interim service preferences of parents waiting for children's mental health treatment: a discrete choice conjoint experiment. J Abnorm Child Psychol 2013 Aug;41(6):865-877. [doi: 10.1007/s10802-013-9728-x] [Medline: 23435482]
39. Chu JT, Wadham A, Jiang Y, Whittaker R, Stasiak K, Shepherd M, et al. Effect of MyTeen SMS-based mobile intervention for parents of adolescents: a randomized clinical trial. JAMA Netw Open 2019 Sep 04;2(9):e1911120 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.11120] [Medline: 31509210]

40. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
41. Data extraction forms. Cochrane. URL: https://dplp.cochrane.org/data-extraction-forms [accessed 2022-02-01]
42. Gulliver A, Griffiths KM, Christensen H, Brewer JL. A systematic review of help-seeking interventions for depression, anxiety and general psychological distress. BMC Psychiatry 2012 Jul 16;12:81 [FREE Full text] [doi: 10.1186/1471-244X-12-81] [Medline: 22799879]
43. Montoya A, Hernández S, Massana MP, Herreros O, Garcia-Giral M, Cardo E, et al. Evaluating Internet information on attention-deficit/hyperactivity disorder (ADHD) treatment: parent and expert perspectives. Educ Health (Abingdon) 2013;26(1):48-53 [FREE Full text] [doi: 10.4103/1357-6283.112801] [Medline: 23823673]
44. Ossebaard HC, van Gemert-Pijnen JE, Sorbi MJ, Seydel ER. A study of a Dutch online decision aid for parents of children with ADHD. J Telemed Telecare 2010 Jan 19;16(1):15-19. [doi: 10.1258/jtt.2009.001006] [Medline: 20086262]
45. Ryan GS, Haroon M, Melvin G. Evaluation of an educational website for parents of children with ADHD. Int J Med Inform 2015 Nov;84(11):974-981. [doi: 10.1016/j.ijmedinf.2015.07.008] [Medline: 26265480]
46. Sapru I, Khalid-Khan S, Choi E, Alavi N, Patel A, Sutton C, et al. Effectiveness of online versus live multi-family psychoeducation group therapy for children and adolescents with mood or anxiety disorders: a pilot study. Int J Adolesc Med Health 2016 Oct 14;30(4):/j/ijamh.2018.30.issue-4/ijamh-2016-0069/ijamh-2016-0069.xml. [doi: 10.1515/ijamh-2016-0069] [Medline: 27740923]
47. Sourander A, McGrath PJ, Ristkari T, Cunningham C, Huttunen J, Hinkka-Yli-Salomäki S, et al. Two-year follow-up of internet and telephone assisted parent training for disruptive behavior at age 4. J Am Acad Child Adolesc Psychiatry 2018 Sep;57(9):658-68.e1. [doi: 10.1016/j.jaac.2018.07.001] [Medline: 30196869]
48. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016 Oct 12;355:i4919 [FREE Full text] [doi: 10.1136/bmj.i4919] [Medline: 27733354]
49. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011 Oct 18;343:d5928 [FREE Full text] [doi: 10.1136/bmj.d5928] [Medline: 22008217]
50. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health 1999 Feb;53(2):105-111 [FREE Full text] [doi: 10.1136/jech.53.2.105] [Medline: 10396471]
51. Chapter 1: overview of intervention mapping. In: Planning Health Promotion Programs: An Intervention Mapping Approach. San Francisco: John Wiley & Sons; 2011.
52. Wei Y, McGrath PJ, Hayden J, Kutcher S. Mental health literacy measures evaluating knowledge, attitudes and help-seeking: a scoping review. BMC Psychiatry 2015 Nov 17;15:291 [FREE Full text] [doi: 10.1186/s12888-015-0681-9] [Medline: 26576680]

Abbreviations

ADHD: attention-deficit/hyperactivity disorder
DHI: digital health intervention
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by A Mavragani; submitted 15.03.21; peer-reviewed by S Chokshi, W Sui; comments to author 30.05.21; revised version received 20.09.21; accepted 10.12.21; published 10.02.22.

Please cite as:
Peyton D, Goods M, Hiscock H
The Effect of Digital Health Interventions on Parents' Mental Health Literacy and Help Seeking for Their Child's Mental Health Problem: Systematic Review
J Med Internet Res 2022;24(2):e28771
URL: https://www.jmir.org/2022/2/e28771
doi:10.2196/28771
PMID:35142623

Review

Digital Health Interventions for Weight Management in Children and Adolescents: Systematic Review and Meta-analysis

Matina Kouvari[1,2], PhD; Melina Karipidou[1], MSc; Thomas Tsiampalis[1], MSc; Eirini Mamalaki[1], PhD; Dimitrios Poulimeneas[1], PhD; Eirini Bathrellou[1], MSc; Demosthenes Panagiotakos[1,2], DrMedSci; Mary Yannakoulia[1], PhD

[1]Department of Nutrition and Dietetics, School of Health Science and Education, Harokopio University, Athens, Greece

[2]Faculty of Health, University of Canberra, Canberra, Australia

Corresponding Author:
Mary Yannakoulia, PhD
Department of Nutrition and Dietetics
School of Health Science and Education
Harokopio University
Eleftheriou Venizelou 70
Athens, 17671
Greece
Phone: 30 6944362633
Email: myianna@hua.gr

Abstract

Background: Recent meta-analyses suggest the use of technology-based interventions as a treatment option for obesity in adulthood. Similar meta-analytic approaches for children are scarce.

Objective: The aim of this meta-analysis is to examine the effect of technology-based interventions on overweight and obesity treatment in children and adolescents.

Methods: A systematic literature search was performed using MEDLINE (PubMed), Scopus, and Cochrane Library for randomized clinical trials to identify interventional studies published between January 2000 and February 2021.

Results: In total, 9 manuscripts from 8 clinical trials of 582 children or adolescents were considered eligible. BMI, BMI z-score, and other BMI-related baseline metrics during and after intervention were considered as primary outcomes. In 7 of 8 studies, a technology-based intervention was applied in addition to conventional care. Of the 8 studies, 6 studies were conducted in the United States, 1 in Australia, and 1 in northwestern Europe. In total, 5 studies included adolescents, whereas the rest addressed children aged 9 to 12 years. Intervention duration ranged from 3 to 24 months. Significant differences between groups in BMI metric changes were reported by 5 of the 8 studies. Pooled analysis revealed an overall significant decrease in BMI metrics in the intervention group (standardized mean difference –0.61, 95% CI –1.10 to –0.13; *P*=.01). Subgroup analysis revealed that significance was lost in case of no parental involvement (standardized mean difference –0.36, 95% CI –0.83 to 0.11; *P*=.14). The small number of clinical trials found, the varying study quality, and the study heterogeneity are some limitations of this review.

Conclusions: The studies reported herein describe functional and acceptable technology-based approaches, in addition to conventional treatments, to enhance weight loss in young populations.

(J Med Internet Res 2022;24(2):e30675) doi:10.2196/30675

KEYWORDS

childhood obesity; eHealth; mHealth; digital health; youth; mobile phone

Introduction

Background

Excess weight in childhood and adolescence has remained one of the most important global public health challenges since emerging as a concern several decades ago [1]. The urgent need to reverse the course of childhood obesity has led to significant growth in research regarding the efficacy of childhood obesity interventions [2]. Various interventions have been tested so far, from school-based interventions to comprehensive behavioral programs with multiple components, delivered by a multidisciplinary team [3,4]. Such models of treatment—even when effective—are often inconvenient, burdensome, and

inaccessible in some cases. New computer- or mobile-assisted information and communication tools can provide useful means to develop smart digital health interventions that could tackle childhood obesity [5,6]. Data collected through internet-linked systems, electronic health records capturing clinical or demographic information, and sensors or smartphones tracking dietary behaviors provide the opportunity to generate useful knowledge regarding users' health, behavior, and progress [7].

A previous meta-analysis with 83 randomized clinical trials (RCTs) has suggested the use of technology-based interventions as a treatment option for obesity in adulthood with potential benefits in weight loss [8]. For children and adolescents, there is only 1 meta-analysis on eHealth overweight and obesity interventions, where parents or caregivers were the agents of change [9]. The meta-analysis included interventions, such as behavioral websites with nutrition information, interactive voice response sessions, or telemedicine via videoconferencing. The fact that most of the eligible technological facilities lacked an interaction with users and the self-monitoring component is considered a limitation. Other meta-analyses examined the effect of web-based or mobile-based interventions on children commenting on modifications in obesogenic behaviors, such as sedentary lifestyle or unhealthy nutritional habits and not on core outcomes such as BMI [10,11]. Therefore, this systematic review and meta-analysis examines the effect of technology-based interventions on overweight and obesity treatment in childhood and adolescence.

Objectives

The objective of this meta-analysis is to determine whether such interventions, delivered mostly on top of conventional care, could be more effective in improving the weight status of children or adolescents with overweight or obesity compared with conventional care or no care. The research hypothesis in this study is that technology-based interventions are effective in weight management and in case of direct comparison with conventional care, at least equivalent to conventional care.

Methods

Search Strategy

Following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2009 guidelines, a computer-assisted systematic literature search (not a registered protocol) was performed by 2 independent researchers (M Kouvari and M Karipidou) using MEDLINE (PubMed), Scopus, and the Cochrane Library for RCTs examining the effect of technology-based versus conventional interventions on weight management of children and adolescents with excess weight. The search strategy was mainly based on Medical Subject Headings terms as follows: *(obesity* OR *overweight* OR *body mass index* OR *weight* OR *diet* OR *nutrition)* AND *(mobile health* OR *ehealth* OR *mhealth* OR *mobile technology* OR *Internet* OR *cellular phone* OR *cellular phones* OR *smartphone* OR *telecommunications* OR *mobile applications* OR *web-based* OR *mobile apps* OR *portable electronic app* OR *portable software app* OR *text message* OR *SMS* OR *short message service* OR *portable game* OR *computers, handheld* OR *PDA* OR *personal digital assistant* OR *social media* OR *social media health* OR *Twitter* OR *tweets* OR *Facebook* OR *Instagram* OR *mobile fitness apps* OR *online social networking* OR *virtual reality* OR *avatars* OR *online gaming* OR *video games)* AND *(pediatric* OR *child* OR *adolescent* OR *youth)* AND *(clinical trial* OR *pilot study* OR *randomized controlled clinical trial)*. The search was limited to publications in English from January 1, 2000, to February 1, 2021. Reference lists of retrieved articles were also considered when these were relevant to the issue examined yet not allocated in the basic search. The relevance of the studies was assessed using a hierarchical approach based on the title, abstract, and full manuscript.

Titles and abstracts of the identified studies were independently screened by 2 researchers (M Kouvari and M Karipidou), and duplicates were removed. Full-text copies of papers were assessed for eligibility (M Kouvari and M Karipidou), with any disagreements resolved by a third researcher (EB). Data for each included study were extracted by 1 researcher (M Kouvari) and cleaned and checked by another (M Karipidou). The 2 researchers (M Kouvari and M Karipidou) extracted data using a standardized extraction form to ensure that it adequately captured trial data. For papers in which additional information was required, the corresponding authors were contacted via email.

Selection Criteria

Studies were selected based on the inclusion and exclusion criteria presented in Textbox 1.

Quality Assessment of Selected Studies

The quality assessment of the selected validation studies was independently implemented by 2 researchers (M Kouvari and M Karipidou) using the Consolidated Standards of Reporting Trials statement [12]. Any differences were discussed, and a decision was made by consensus.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Study design: controlled clinical trials with at least one arm with a technology-based intervention controlled by a second arm with a conventional care intervention or without any intervention
- Sample: children and adolescents with overweight or obesity (defined through BMI or validated growth charts) aged ≤18 years
- Intervention: technology-based intervention for children or adolescents with or without parents' or families' support
- Outcome: BMI, BMI *z*-score, and other BMI-related metrics (eg, BMI-SD score) at baseline, during the intervention, and at the postintervention phase were considered as the primary measurements for this meta-analysis

Exclusion criteria

- Review articles
- Letters to editors
- Editorials
- Articles based on studies with adults
- Articles providing only feasibility or acceptance level of the applied technology-based interventions or outcomes related only to obesogenic behaviors
- Articles in which the technology-based intervention was applied only to parents
- Articles in which the control group included the use of technology
- Articles in which the technology-based intervention was not interactive with the user, for example, telemedicine or it had only an informative character, for example, a website
- Articles with inadequate statistical information

Effect Size Measurements

The outcome of interest in this meta-analysis was the difference between the web-based intervention and the control group with regard to the potential changes from cumulative frequency distribution in BMI or BMI z-score or BMI-SD score. Studies that reported BMI-related metric results as change scores or baseline and final values; SD, SE, or CIs; and number of participants in each intervention group were included in the meta-analysis. The mean change was calculated where required, and SDs were calculated from SE or 95% CI where SD was not reported [13]. Finally, missing SDs of the changes from baseline were calculated using an imputed correlation coefficient [13].

Data Analysis

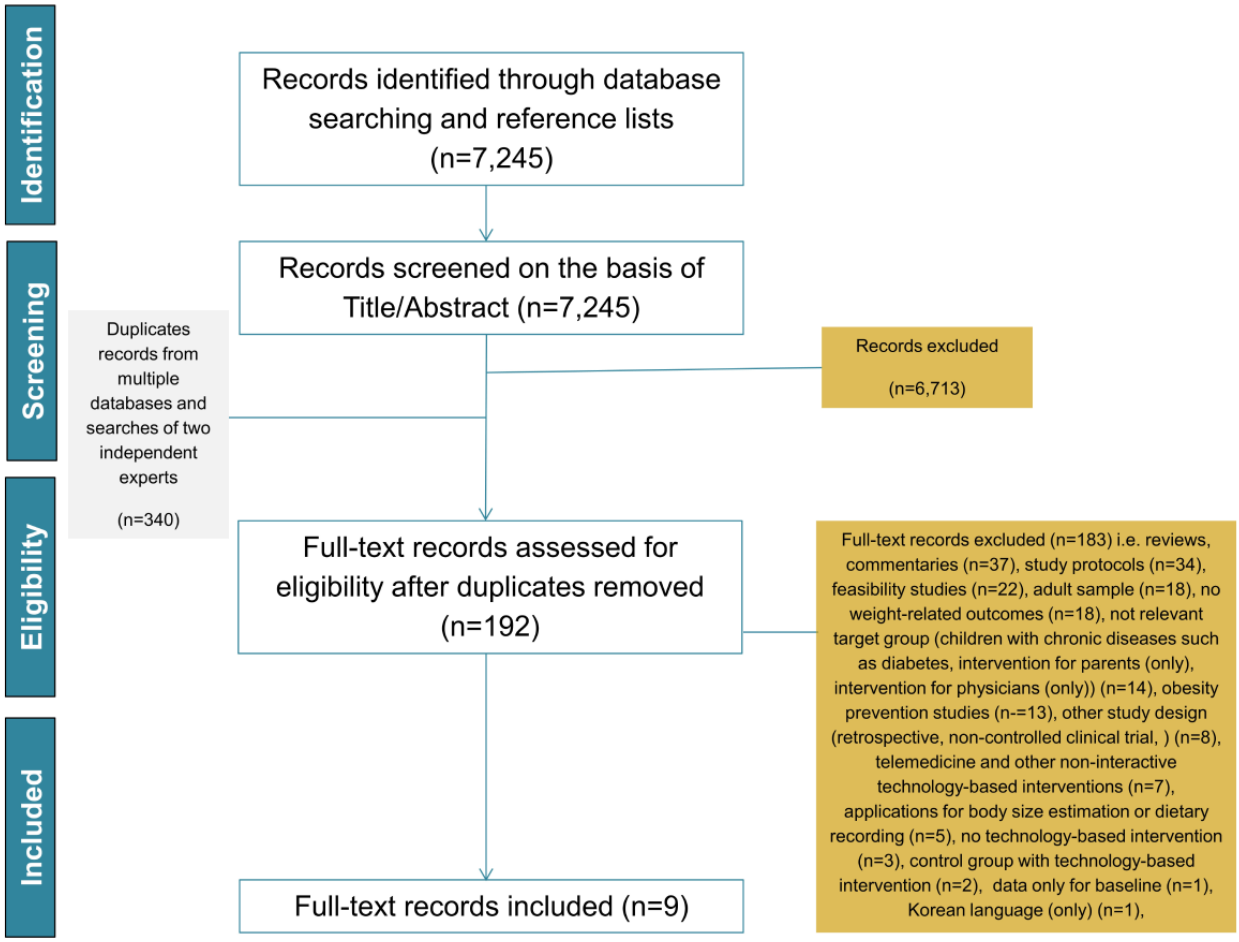
Standardized mean difference (SMD) was used to enable the inclusion of BMI-related metrics in the same meta-analysis. In a study that reported >1 BMI metric, BMI was used. Pooled values of SMDs between the technology-based intervention and the control group and 95% CIs as the recommended summary statistics of the effect size were calculated using either a fixed or random effects model. The fixed effects model was used when sample heterogeneity was <50%, and the random effects model was used when heterogeneity was >50%. Heterogeneity assessed the null hypothesis that all studies evaluated the same effect and was evaluated using the chi-square test. Inconsistency (I_2) was calculated to quantify the total variation consistent with interstudy heterogeneity, ranging from 0% to 100%. A *P* value of <.10 for the chi-square test and I_2 >50% reflected a significant heterogeneity [14]. Estimates of the effect size measures were weighted by the inverse of their variances. The random effects model (DerSimonian and Laid method) was used in the presence of heterogeneity. In contrast, fixed effects models were used to calculate effect size estimates for studies that lacked heterogeneity. Subgroup analysis of prespecified groupings was performed for the following study characteristics: duration of follow-up (3-24 months), parental involvement, and type of intervention (web based vs mobile based and others). In subgroup analyses, only the last follow-up values were considered. In studies with multiple follow-ups, only the last follow-up time was considered for the estimation of the overall effect size. Possible publication bias was assessed using a contour-enhanced funnel plot of each trial's effect size against the SE. Funnel plot asymmetry was evaluated using the Begg and Egger tests [15]. Stata software, version 14 (StataCorp LLC) was used for all statistical analyses.

Results

Flow of Included Studies

A literature search flow diagram is presented in Figure 1. Initially, 7245 papers were retrieved and selected for evaluation. Then, 6713 manuscripts were removed based on their titles and abstracts as they were irrelevant to the scope of this work, accompanied by 340 duplicate records from multiple databases and searches that were also excluded. Among the rest (n=192), 9 manuscripts from 8 studies (ie, 2 separate articles were published based on 1 study regarding 2 follow-up periods) were considered relevant; 183 manuscripts were excluded, as they did not meet the inclusion criteria of this systematic review.

Figure 1. Flow diagram describing the literature review process.



General Characteristics of the Selected Clinical Trials

The characteristics of the eligible clinical trials for this meta-analysis are presented in Multimedia Appendix 1 [16-24]. In total, 582 children and adolescents participated in the selected 8 studies with a range of cultural or ethnic groups, including African American, Chinese American, White, and others. Of the 8 studies, 6 (75%) studies were conducted in the United States [16-22], 1 (13%) in Australia [23], and 1 (13%) in northwestern Europe (Netherlands) [24]. Moreover, 75% (6/8) of studies were conducted within the last decade [16-19,23,24], whereas the remaining 25% (2/8) of studies were conducted earlier [20-22]. Most of the selected studies addressed adolescents [16,17,20-23], whereas the rest had children aged 9-12 years as the target group [18,19,24]. The length of interventions ranged from 3 months [16,19,24] to 4 months [20], 6 months [17,18], and 24 months [21-23]. All studies were 2-arm controlled clinical trials, in which technology-based interventions were controlled for 1 conventional care intervention, except for 2 studies in which no intervention was implemented in the control group [16,19,24].

Description of the Technology-Based Interventions

Of the 8 studies, 4 (50%) examined the effect of a mobile health (mHealth) intervention with or without sensors [16,17,19,24], 3 (38%) studies used a web-based intervention [18,20-22], and 1 (13%) study used an SMS text messaging intervention accompanied by telemedicine [23]. Focusing on mobile-based interventions, they also addressed nutrition-related issues and unhealthy dietary behaviors [16,17,19,24], whereas in 38% (3/8) of studies, physical activity and screen time were also taken into account [16,19,24]. In 1 study with web-based interventions, participants were enhanced to increase their physical activity level via a gamification method [18]. The other 2 web-based interventions, following a family-oriented approach, provided nutrition education accompanied by physical activity tips and counseling regarding healthy body image [20-22]. The SMS text messaging intervention along with telemedicine and group sessions focused on weight loss and weight loss maintenance, covering issues from nutrition and physical activity to body image and psychological well-being [23]. The level of parental involvement varied among the selected studies. In 6 (75%) of the 8 studies, there was participation of parents in the intervention group [18-24] accompanied by a similar participation of parents in the control group, with the exception of 2 (25%) studies [18,19]. In 7 (88%) out of 8 interventions, a hybrid approach was followed, which means that the technological tools—of any kind—were examined as supportive of conventional care treatment [17,19-24]. In 7 (88%) of the 8 studies, there was support from health care practitioners, such as dietitians, physicians, pediatricians, and psychologists [17-24]. Participants in the intervention group (children or adolescents alone or with their parents) attended weekly, biweekly, or monthly face-to-face sessions with health professionals [17,19-24] or videoconferences [18]. These sessions included goal setting, motivation techniques,

individualized feedback based on the technology-based dietary or physical activity records, and enhancement to use the digital tools provided.

Primary and Secondary Outcomes of the Selected Clinical Trials

Different measures of weight status and adiposity were used in the selected studies, with most of them using multiple measures. In total, of the 8 studies, 5 (63%) used BMI z-score [17-20,23], 3 (38%) used BMI [16,19,21,22], 3 (38%) used BMI percentiles [17,19,21,22], 2 (25%) used body fat [18,21,22], 1 (13%) used waist-to-hip ratio [23], and 1 (13%) used BMI-SD score [24]. Other metrics included modifications in obesogenic behaviors, such as dietary habits [16,18-20,22-24], physical activity habits and/or screen time [16,18,19,22,23], and physical examination or biochemical metrics [18,23]. All studies included psychological and self-efficacy metrics related to diet, physical activity, well-being, or healthy body image. With the exception of 25% (2/8) of studies [16,23], the remaining studies provided information on participants' satisfaction and compliance with the technology-based intervention.

Risk of Bias Within Selected Studies

The results of the risk of bias assessment for all included studies are summarized in Table 1. The selected eligible studies were of moderate quality, meeting on average, approximately 6 out of the 9 quality criteria. In particular, all studies except 1 had a well-documented randomization process [17]. In all studies, except for 1 study [22], the baseline characteristics were presented. All studies used a valid method to assess the main outcome of interest, that is, BMI, whereas only 4 (44%) out of 9 studies reported blinded assessment of the outcome of interest [16,18,20,23]. All studies except 1 [23] met the dropout rate cut-off points (ie, ≤20% for <6 months and ≤30% for ≥6 months). Regarding the quality of statistical analysis, on average, the selected studies met 2 out of 9 criteria. Specifically, all studies except 3 used intention-to-treat analysis [17,20,24]; all studies except 2 reported adequate statistical power [17,19], whereas only 4 studies provided adjusted differences between groups [16-18,23].

Table 1. Quality assessment of the eligible clinical trials (9 manuscripts and 8 studies)[a].

Characteristics	Study								
	Chen et al [16]	Vidmar et al [17]	Staiano et al [18]	Wright et al [19]	Nguyen et al [23]	de Niet et al [24]	Doyle et al [20]	Williamson et al [22]	Williamson et al [21]
Study design									
Randomization described and conducted	✓		✓	✓	✓	✓	✓	✓	✓
Baseline characteristics by group	✓	✓	✓	✓	✓	✓	✓		✓
Outcome assessment									
Valid measurement of BMI	✓	✓	✓	✓	✓	✓	✓	✓	✓
Blinded outcome assessment	✓		✓		✓		✓		
Dropout rate									
≤20% for <6 months and ≤30% for ≥6 months	✓	✓	✓	✓		✓	✓	✓	✓
Statistical analysis									
Intention to treat for BMI outcomes	✓		✓	✓	✓			✓	✓
Covariates accounted for in analysis	✓		✓		✓			✓	✓
Power calculation reported and power adequate	✓		✓		✓	✓	✓	✓	
Summary results, adjusted difference between groups, and CI	✓	✓	✓		✓				
Scoring									
Score in study design (range 0-2)	2	1	2	2	2	2	2	1	2
Score in outcome assessment (range 0-2)	2	1	2	1	2	1	2	1	1
Score in dropout rate (range 0-1)	1	1	1	1	0	1	1	1	1
Score in statistical analysis (range 0-4)	4	1	4	1	4	1	1	3	2
Total score (range 0-9)	9	4	9	5	8	5	6	6	6

[a]Quality assessment was performed based on the Consolidated Standards of Reporting Trials statement.

Separate Outcomes of Selected Studies

Overview

The separate outcomes of the eligible clinical trials are summarized in Multimedia Appendix 2 [16-24].

Weight and Adiposity Outcomes

Of the 8, 5 (63%) studies reported significant differences between groups in BMI metrics from baseline to the end of intervention [18,20,21,23,24]. The intervention duration of these studies was >6 months, and all of them addressed not only children or adolescents but also their parents. Significant reductions in body fat [22] and waist-to-hip ratio [23] were observed in interventions with a 2-year duration.

Diet-Related Outcomes

The 7 studies reporting modifications on dietary intake and behaviors revealed a significant difference between groups with regard to improvement in at least one dietary outcome. In particular, a decrease in consumption of sugar-sweetened beverages [16], lower carbohydrate intake [18], increased fruit consumption [19], decreased meat and fruit juice intake [23], better adherence to a healthier dietary pattern [20,24], and lower consumption of food products with high-fat content [21,22] were observed.

Physical Activity–Related Outcomes

Among the 5 studies that provided input on changes in participants' physical activity level, 1 (20%) study revealed a

significant decrease in screen time [16], whereas the remaining 4 (80%) studies highlighted increases in physical activity level in terms of hours per day or the intensity of exercise [17,20,21,24].

Physical Examination and Biochemical Metrics

In 1 (50%) of the 2 studies with physical examination and biochemical measurements, significant reductions in blood pressure and cholesterol levels were observed [18].

Psychological Health–Related Outcomes

All studies provided input on the effect of technology-based intervention over the control group on participants' psychological health. Of the 8 studies, 7 (88%) studies observed that participants in the intervention group increased their self-efficacy in relation to diet [16,19,20] and physical activity [16,18,20], decreased unhealthy eating behaviors related to dieting or weight or body image [22,23], and ameliorated their self-esteem [23,24].

Usability and Acceptability of the Technology-Based Intervention

Of the 6 studies providing information on the level of compliance of participants with the technology-based intervention, 5 (83%) reported moderate to high levels of usability and acceptability of the technology-based intervention [17-20,24]. Nevertheless, the added value of the technology-based intervention over typical care was not clear considering the similar dropout rates between the 2 groups (dropout rate in the intervention group: 12.9%, range 0%-41%, vs dropout rate in the control group: 12.6%, range 0%-36%; *P*=.96), excluding studies in which the control group had no intervention [18,19]. Among the 6 studies, 1 (17%) revealed that children assigned to the technology-based intervention receiving SMS text messages were less likely to withdraw from the study than children who did not receive this service [24].

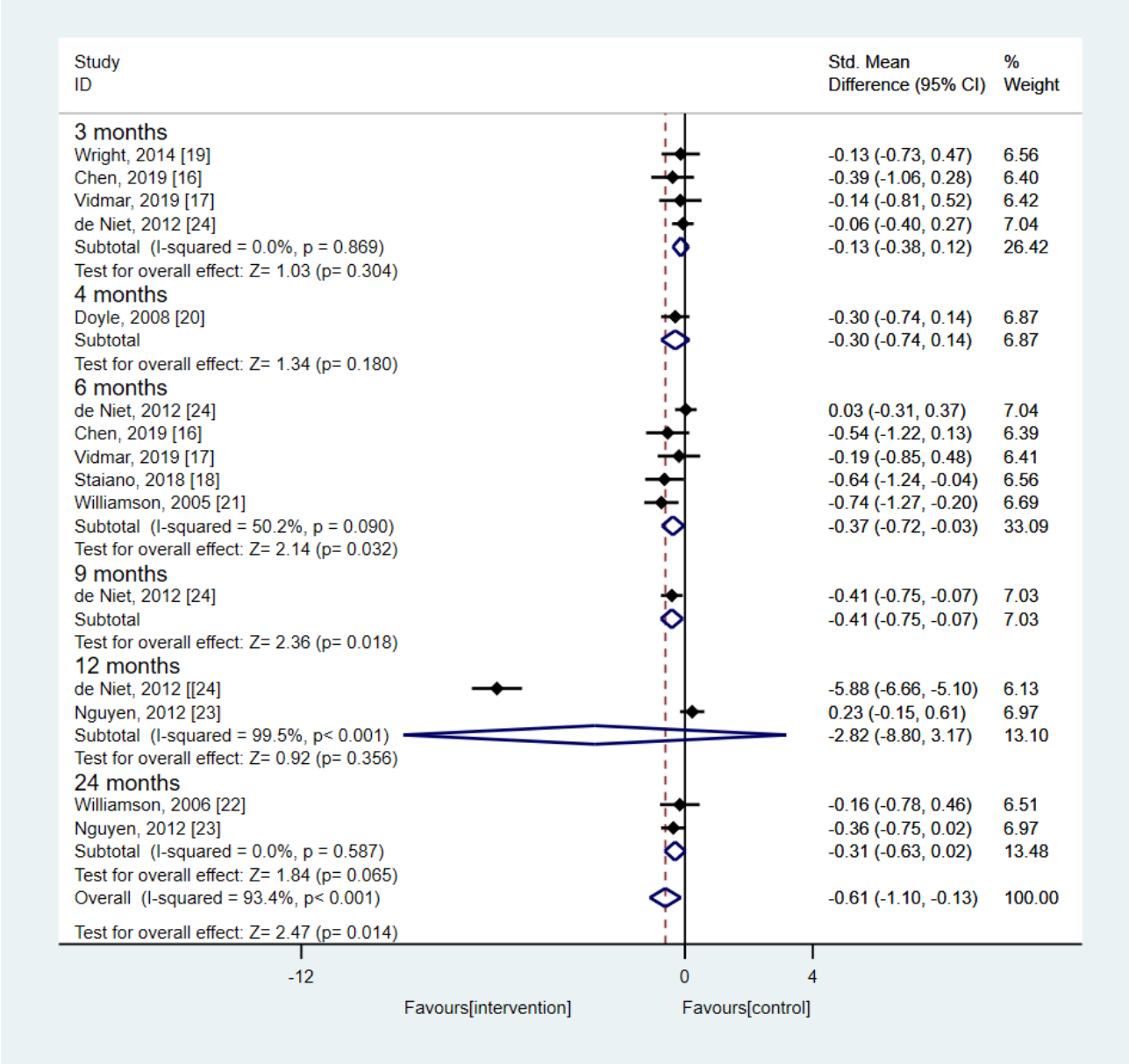
Synthesis of BMI-Related Outcomes

Overview

A meta-analysis was conducted on pooled data from 9 manuscripts (8 studies in total), which compared technology-based intervention groups with control groups. The meta-analysis results are presented in Figures 2-4.

As presented in Figure 2, a significantly higher decrease in the BMI-related metric was observed (SMD –0.61, 95% CI –1.10 to –0.13; *P*=.01). Compared with the other follow-ups, this was more evident after a 6-month follow-up in the technology-based intervention group when compared with the control group (SMD –0.37, 95% CI –0.72 to –0.03; *P*=.03), whereas a favorable effect of the technology-based interventions was also found after 24 months; however, statistical significance was not reached (SMD –0.31, 95% CI –0.63 to 0.02; *P*=.07).

Figure 2. Results from the random effects meta-analysis concerning the effect of the technology-based interventions on BMI-related metrics according to the study follow-ups. In case of studies with multiple follow-ups, only the last follow-up time was considered for the estimation of the overall effect size.

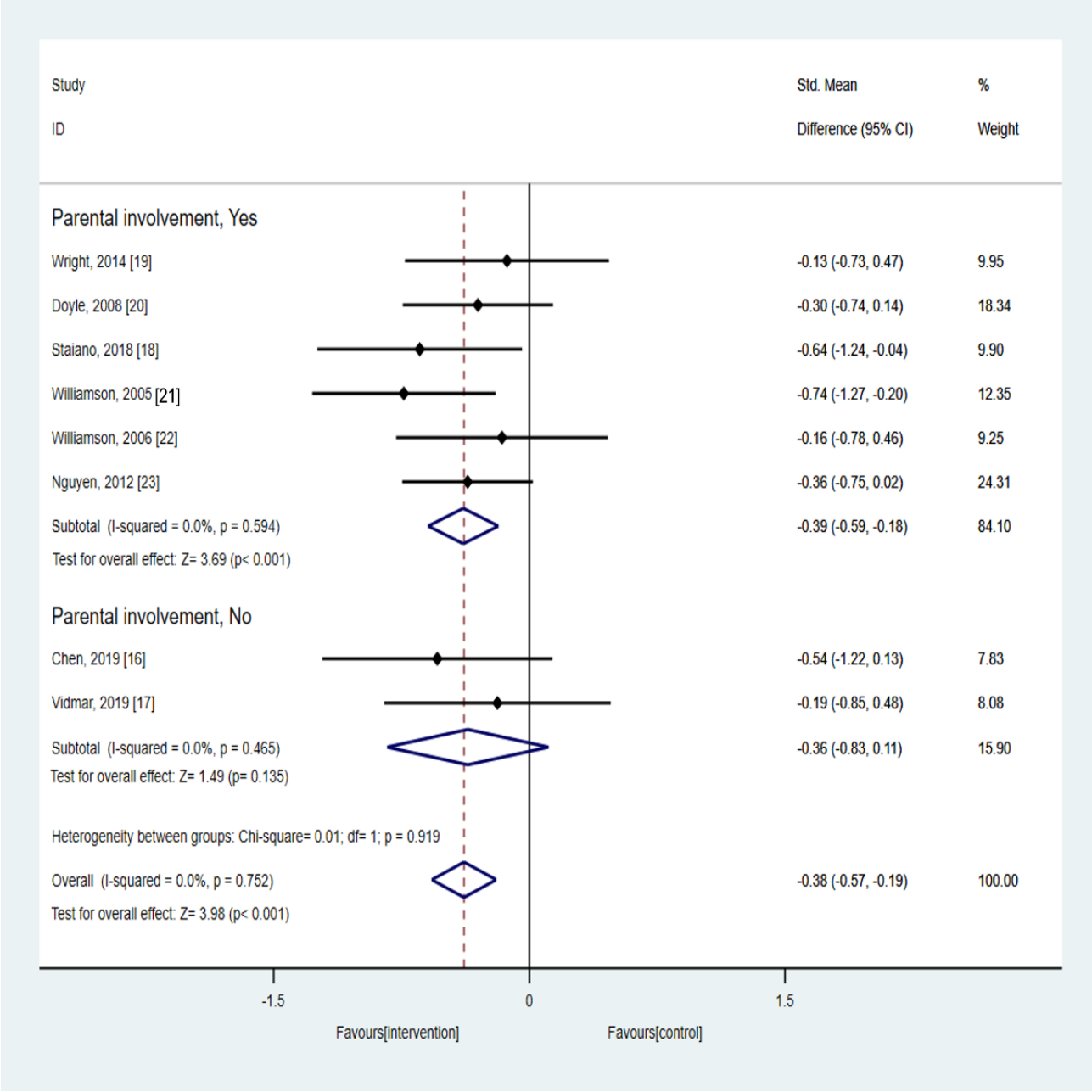


Sensitivity Analysis

Of the 8 studies, 2 (25%) had a control group without any intervention. Hence, we repeated the aforementioned analysis, excluding these 2 studies. The overall outcome remained significant (SMD –0.65, 95% CI –1.20 to –0.10; *P*=.02), whereas the 6-month outcome remained marginally significant (SMD –0.32, 95% CI –0.71 to 0.07; *P*=.10).

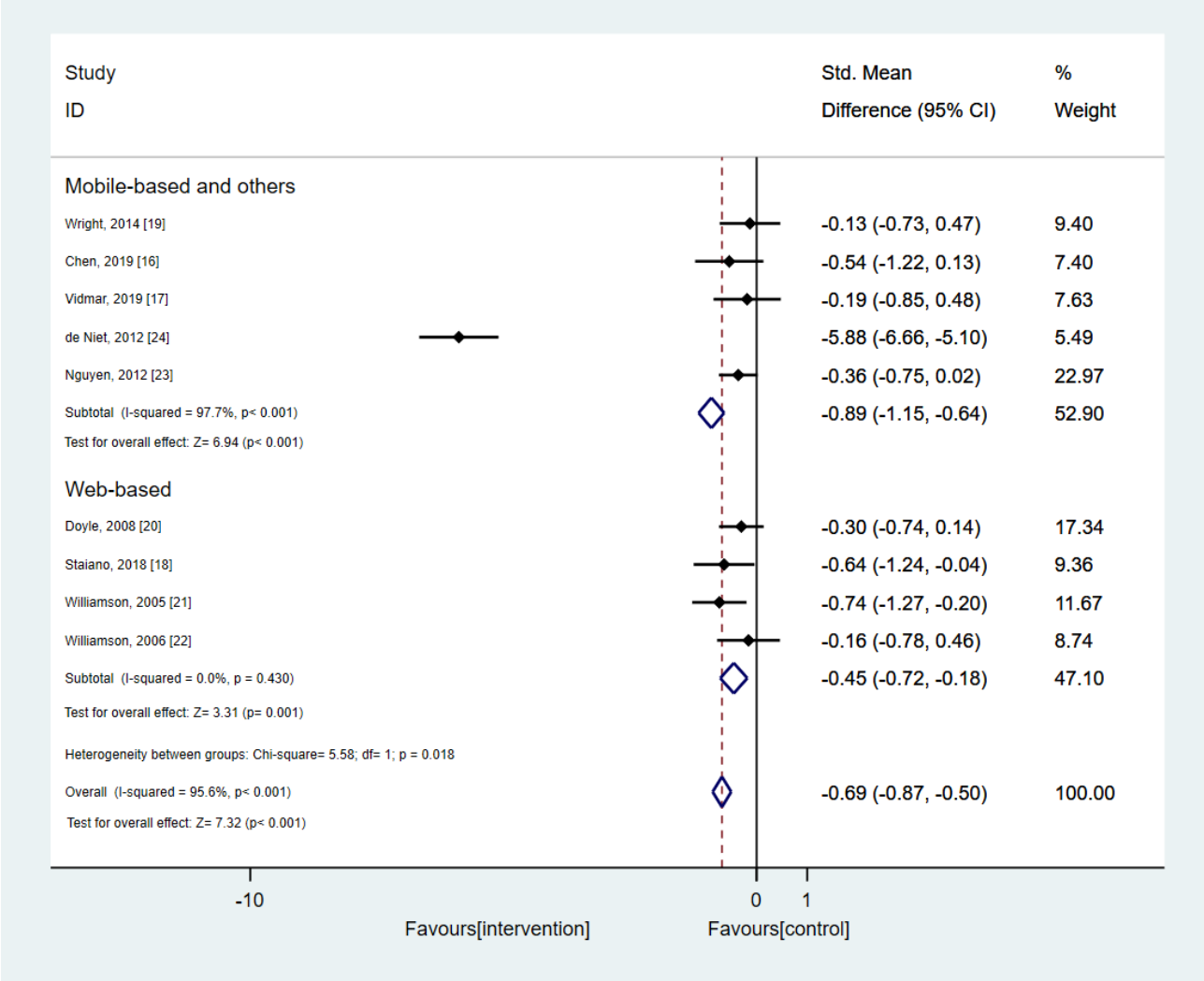
A subgroup analysis was conducted based on parental involvement, and the results are shown in Figure 3. The meta-analysis revealed a significantly higher decrease in the BMI-related metric in the technology-based intervention group than in the control group only in case of parental involvement (SMD –0.39, 95% CI –0.59 to –0.18; *P*<.001). In the case of no parental involvement, no significant difference between groups was observed (SMD –0.36, 95% CI –0.83 to 0.11; *P*=.14).

Figure 3. Results from the subgroup analysis according to the parental involvement of the technology-based intervention concerning its effect on BMI-related metrics. In case of studies with multiple follow-ups, only the last follow-up time was considered for the estimation of the overall effect size.



Another subgroup analysis performed in this study was related to the type of technology-based intervention used. The results are shown in Figure 4. Interventions were grouped as web-based, mobile-based, and others. A statistically significant decrease in the BMI-related metric in the intervention group compared with that in the control group was observed both in the case of mobile-based and other interventions (SMD –0.89, 95% CI –1.15 to –0.64; *P*<.001) as well as in the case of the web-based interventions (SMD –0.45, 95% CI –0.72 to –0.18; *P*=.001).

Figure 4. Results from the subgroup analysis according to the type of the intervention concerning its effect on BMI-related metrics. In case of studies with multiple follow-ups, only the last follow-up time was considered for the estimation of the overall effect size.



Discussion

Principal Findings

This meta-analysis revealed that an intervention that combines conventional care with technological facilities could be an effective method for weight management of children and adolescents with overweight or obesity and is probably more effective than conventional care alone. These observations were more evident in the case of interventions lasting at least six months. The selected studies included eHealth and mHealth technologies, such as interactive web platforms, mobile apps, gaming, and SMS text messaging with or without sensors and were accompanied or not accompanied by other contact forms such as telemedicine, emails, and informative websites. The focus of these technologies was more or less related either exclusively or in combination with improvement in dietary habits, enhancement of physical activity, or the increasing of users' self-monitoring potential. The type of technological means used in each intervention, that is, mobile-based or web-based, did not seem to alter the final outcome. Parental involvement was related to greater outcomes of the intervention, particularly in children; however, it was not possible to isolate the separate contribution of parents to the final outcome.

Strengths and Limitations

To the best of our knowledge, this is the first meta-analysis to examine the effect of technology-based interventions on weight management in childhood and adolescence using strict eligibility criteria, such as the existence of a control—without any kind of technology—group, the exclusion of technology-based interventions without an interactive—with the user—character, and the exclusion of studies providing the effect on weight-related metrics of normal-weight children or adolescents (obesity prevention spectrum). However, several limitations exist, including the restriction to articles published only in English, the small number of clinical trials found with varying study quality, heterogeneity of the studies, inadequacy of the power to detect an outcome in some studies because of the small number of participants, varying aims between studies, and all but 2 studies being conducted in the United States. Finally, no definite conclusion can be drawn on whether the actual difference in weight management was caused by the different techniques used or by the differences in other intervention characteristics (eg, number of sessions).

Comparison With Prior Work

To date, technology-based health interventions for weight management in young people have been reported as a rapidly developing research area; although promising results have been produced, more research is definitely in order [25]. Hitherto, studies have described technologies with the potential to promote healthy behaviors related to nutrition or physical activity in children and adolescents [26,27]. The pooled higher reduction in adiposity metrics observed in this study in favor of technology-based intervention groups were observed with improvements in users' dietary habits, physical activity level, and screen time and with better psychological health regarding body image, self-esteem, or weight- and dieting-related stress. Examining the actual efficiency and added value of technology-based methods over conventional interventions in childhood obesity management remains a challenging research field for many reasons. Most interventional studies combine eHealth or mHealth interventions with other mostly conventional treatments (hybrid approach), thus making the efficacy of technology-only approaches to affect adiposity outcomes difficult to ascertain [9,26]. In most studies selected for this meta-analysis, the digital behavior change intervention was examined on top of conventional care and matched for conventional-only care, revealing an advantage of the former in weight loss. On the other hand, digital behavior change interventions have not been generated to replace the role of health professionals and the multifaceted treatments required for management of excess weight in childhood and adolescence but rather to support them, showing potential as an additional tool for patient monitoring and designing tailor-made interventions through the selection of more valid information and plausibly weight loss maintenance in the postintervention phase [28,29].

Another issue examined in this study was related to parental involvement in technology-based interventions. Active enrollment of parents was reported in 6 (75%) out of 8 studies [18-24], accompanied by a similar participation in the control group, with the exception of 2 studies [18,19]. Interestingly, 2-8 studies without parental involvement did not achieve significant modifications in adiposity metrics. Although many reasons could be responsible for this nonsignificance, such as the fact that lower combined sample size in the 2 studies could lead to greater CIs and higher *P* values, this finding may imply that the participation of families in childhood obesity management programs remains of high importance even in or especially in interventions with advanced technological means. Currently, interventions that target parents to tackle obesity in early life stages are presented as effective, especially when it comes to preschoolers, that is, children <5 years [30-33]. On the other hand, the studies in this meta-analysis were designed for children >9 years and principally adolescents, which may challenge the level of parental involvement. Preschool-aged children are rarely targeted in such technology-based interventions. The MINISTOP RCT is probably the very first study to apply a 6-month technology-based weight loss intervention to children <5 years, using their parents as the primary target group. Although no significant differences in adiposity metrics between the control and intervention groups were observed, children and parents assigned to the technology-based group seemed to significantly ameliorate their nutrition and physical activity habits [34]. Considering the fact that technological approaches and parental involvement are usually presented as effective practices to tackle excess weight in childhood, their combination into 1 tailor-made weight loss intervention may result in multiple positive outcomes.

The studies in this meta-analysis provided input on the usability and acceptability of technology-based interventions. Most of them reported a moderate to high level of adherence to the intervention using different criteria and metrics, such as the level of user enjoyment, the frequency of app use, or the number of SMS text messages received [17-20,24]. Nevertheless, based on the dropout rates, no significant differences were observed between the intervention and control groups in most cases. This evidence regarding the usability of technology-based interventions is based largely on the use of SMS text messaging using a mobile- or web-based approach. However, the latest technological advances include the emergence of smartphone apps [29], interactive platforms [35], and exergaming [36]. Such facilities have increased in popularity, offering a unique opportunity to implement large-scale obesity treatment interventions in youth [37]. The current orientation for improving adherence to treatments in pediatrics focuses on motivation, problem-solving skills, and reduction of posttreatment influence, resorting to several novel youth-friendly technological approaches [26]. For instance, studies have described the best placement and accuracy of mobile devices and sensors to record dietary intake or physical activity and ways to lessen user burden [26]. Reward-type incentives, provision of social connections and multiplayer capabilities, short- and long-term motivational techniques, and personalized feedback are also suggested as means to enhance user acceptability, efficiency of the intervention, and probably maintenance of positive outcomes even in the postintervention phase [26]. Focusing on gamification, many video games have been created with the aim to modify children's or adolescents' dietary habits or physical activity status, such as *Let's Move! (To move!)*, *Counting Carbohydrates with Lenny*, *LeapBand*, or *Zamzee*, where users interact with virtual characters that—creating a fascinating environment—enhance them to complete a series of relevant activities and challenges [38]. Finally, early involvement of key stakeholders in the intervention development stage seems to be detrimental for the delivery of a technological tool that will be well-accepted by the target group—even more when it comes to younger ages [39].

Conclusions

Studies reported herein describe functional and acceptable technology-based approaches, on top of conventional care, to enhance weight loss in overweight or obese children and adolescents through the promotion of a healthy lifestyle and improvement of users' well-being. However, the large heterogeneity in study designs, settings, intervention components, and outcomes probably eliminates the strength of this conclusion. Finally, this field is advancing so quickly that the technology used is often no longer state of the art; interventions that use the full range of novel technologies, such

as ubiquitous sensing and real-time feedback, are currently being developed and pilot tested. Therefore, similar meta-analytic approaches should be repeated on a regular basis.

Acknowledgments

This work has received funding from the European Union's Horizon 2020 Research and Innovation Programme through the NUTRISHIELD project under grant agreement 818110. This paper reflects only the authors' views; the European Union is not liable for any use that may be made of the information contained therein.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Characteristics of the eligible clinical trials of the meta-analysis (9 manuscripts and 8 studies).
[DOCX File , 26 KB - jmir_v24i2e30675_app1.docx]

Multimedia Appendix 2
Primary and secondary outcomes of the eligible clinical trials of the meta-analysis (9 manuscripts and 8 studies).
[DOCX File , 24 KB - jmir_v24i2e30675_app2.docx]

References

1. Lobstein T, Jackson-Leach R, Moodie ML, Hall KD, Gortmaker SL, Swinburn BA, et al. Child and adolescent obesity: part of a bigger picture. Lancet 2015 Jun;385(9986):2510-2520. [doi: 10.1016/s0140-6736(14)61746-3]
2. Di Cesare M, Sorić M, Bovet P, Miranda JJ, Bhutta Z, Stevens GA, et al. The epidemiological burden of obesity in childhood: a worldwide epidemic requiring urgent action. BMC Med 2019 Nov 25;17(1):212 [FREE Full text] [doi: 10.1186/s12916-019-1449-8] [Medline: 31760948]
3. Scott-Sheldon LA, Hedges LV, Cyr C, Young-Hyman D, Khan LK, Magnus M, et al. Childhood obesity evidence base project: a systematic review and meta-analysis of a new taxonomy of intervention components to improve weight status in children 2-5 years of age, 2005-2019. Child Obes 2020 Sep 01;16(S2):221-248 [FREE Full text] [doi: 10.1089/chi.2020.0139] [Medline: 32936038]
4. Wang Y, Cai L, Wu Y, Wilson RF, Weston C, Fawole O, et al. What childhood obesity prevention programmes work? A systematic review and meta-analysis. Obes Rev 2015 Jul;16(7):547-565 [FREE Full text] [doi: 10.1111/obr.12277] [Medline: 25893796]
5. Lau PW, Lau EY, Wong DP, Ransdell L. A systematic review of information and communication technology-based interventions for promoting physical activity behavior change in children and adolescents. J Med Internet Res 2011;13(3):e48 [FREE Full text] [doi: 10.2196/jmir.1533] [Medline: 21749967]
6. Tate EB, Spruijt-Metz D, O'Reilly G, Jordan-Marsh M, Gotsis M, Pentz MA, et al. mHealth approaches to child obesity prevention: successes, unique challenges, and next directions. Transl Behav Med 2013 Dec;3(4):406-415 [FREE Full text] [doi: 10.1007/s13142-013-0222-3] [Medline: 24294329]
7. Smith AJ, Skow A, Bodurtha J, Kinra S. Health information technology in screening and treatment of child obesity: a systematic review. Pediatrics 2013 Mar;131(3):894-902. [doi: 10.1542/peds.2012-2011] [Medline: 23382447]
8. Hutchesson MJ, Rollo ME, Krukowski R, Ells L, Harvey J, Morgan PJ, et al. eHealth interventions for the prevention and treatment of overweight and obesity in adults: a systematic review with meta-analysis. Obes Rev 2015 May;16(5):376-392. [doi: 10.1111/obr.12268] [Medline: 25753009]
9. Hammersley ML, Jones RA, Okely AD. Parent-focused childhood and adolescent overweight and obesity eHealth interventions: a systematic review and meta-analysis. J Med Internet Res 2016 Jul 21;18(7):e203 [FREE Full text] [doi: 10.2196/jmir.5893] [Medline: 27443862]
10. Shin Y, Kim SK, Lee M. Mobile phone interventions to improve adolescents' physical health: a systematic review and meta-analysis. Public Health Nurs 2019 Nov;36(6):787-799. [doi: 10.1111/phn.12655] [Medline: 31456259]
11. Champion KE, Parmenter B, McGowan C, Spring B, Wafford QE, Gardner LA, Health4Life team. Effectiveness of school-based eHealth interventions to prevent multiple lifestyle risk behaviours among adolescents: a systematic review and meta-analysis. Lancet Digit Health 2019 Sep;1(5):206-221 [FREE Full text] [doi: 10.1016/S2589-7500(19)30088-3] [Medline: 33323269]
12. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Ann Intern Med 2010 Jun 01;152(11):726-732 [FREE Full text] [doi: 10.7326/0003-4819-152-11-201006010-00232] [Medline: 20335313]

13. Cumpston M, Li T, Page M, Chandler J, Welch V, Higgins J, et al. Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions. Cochrane Database Syst Rev 2019 Oct 03;10:ED000142. [doi: 10.1002/14651858.ED000142] [Medline: 31643080]
14. Lau J, Ioannidis JP, Schmid CH. Quantitative synthesis in systematic reviews. Ann Intern Med 1997 Nov 01;127(9):820-826. [doi: 10.7326/0003-4819-127-9-199711010-00008] [Medline: 9382404]
15. Irwig L, Macaskill P, Berry G, Glasziou P. Bias in meta-analysis detected by a simple, graphical test. Graphical test is itself biased. Br Med J 1998 Feb 07;316(7129):470-471 [FREE Full text] [Medline: 9492687]
16. Chen J, Guedes CM, Lung AE. Smartphone-based healthy weight management intervention for Chinese American adolescents: short-term efficacy and factors associated with decreased weight. J Adolesc Health 2019 Apr;64(4):443-449. [doi: 10.1016/j.jadohealth.2018.08.022] [Medline: 30409751]
17. Vidmar AP, Pretlow R, Borzutzky C, Wee CP, Fox DS, Fink C, et al. An addiction model-based mobile health weight loss intervention in adolescents with obesity. Pediatr Obes 2019 Feb;14(2):e12464 [FREE Full text] [doi: 10.1111/ijpo.12464] [Medline: 30117309]
18. Staiano AE, Beyl RA, Guan W, Hendrick CA, Hsia DS, Newton RL. Home-based exergaming among children with overweight and obesity: a randomized clinical trial. Pediatr Obes 2018 Nov;13(11):724-733 [FREE Full text] [doi: 10.1111/ijpo.12438] [Medline: 30027607]
19. Wright JA, Phillips BD, Watson BL, Newby PK, Norman GJ, Adams WG. Randomized trial of a family-based, automated, conversational obesity treatment program for underserved populations. Obesity (Silver Spring) 2013 Sep;21(9):369-378 [FREE Full text] [doi: 10.1002/oby.20388] [Medline: 23512915]
20. Doyle AC, Goldschmidt A, Huang C, Winzelberg AJ, Taylor CB, Wilfley DE. Reduction of overweight and eating disorder symptoms via the internet in adolescents: a randomized controlled trial. J Adolesc Health 2008 Aug;43(2):172-179 [FREE Full text] [doi: 10.1016/j.jadohealth.2008.01.011] [Medline: 18639791]
21. Williamson DA, Martin PD, White MA, Newton R, Walden H, York-Crowe E, et al. Efficacy of an internet-based behavioral weight loss program for overweight adolescent African-American girls. Eat Weight Disord 2005 Sep;10(3):193-203. [doi: 10.1007/BF03327547] [Medline: 16277142]
22. Williamson DA, Walden HM, White MA, York-Crowe E, Newton RL, Alfonso A, et al. Two-year internet-based randomized controlled trial for weight loss in African-American girls. Obesity (Silver Spring) 2006 Jul;14(7):1231-1243. [doi: 10.1038/oby.2006.140] [Medline: 16899804]
23. Nguyen B, Shrewsbury VA, O'Connor J, Steinbeck KS, Hill AJ, Shah S, et al. Two-year outcomes of an adjunctive telephone coaching and electronic contact intervention for adolescent weight-loss maintenance: the Loozit randomized controlled trial. Int J Obes (Lond) 2013 Mar;37(3):468-472. [doi: 10.1038/ijo.2012.74] [Medline: 22584456]
24. de NJ, Timman R, Bauer S, van den Akker E, Buijks H, de KC, et al. The effect of a short message service maintenance treatment on body mass index and psychological well-being in overweight and obese children: a randomized controlled trial. Pediatr Obes 2012 Jun;7(3):205-219. [doi: 10.1111/j.2047-6310.2012.00048.x] [Medline: 22492669]
25. Tully L, Burls A, Sorensen J, El-Moslemany R, O'Malley G. Mobile health for pediatric weight management: systematic scoping review. JMIR Mhealth Uhealth 2020 Jun 03;8(6):e16214 [FREE Full text] [doi: 10.2196/16214] [Medline: 32490849]
26. Turner T, Spruijt-Metz D, Wen CK, Hingle MD. Prevention and treatment of pediatric obesity using mobile and wireless technologies: a systematic review. Pediatr Obes 2015 Dec;10(6):403-409. [doi: 10.1111/ijpo.12002] [Medline: 25641770]
27. Lee J, Piao M, Byun A, Kim J. A systematic review and meta-analysis of intervention for pediatric obesity using mobile technology. Stud Health Technol Inform 2016;225:491-494. [Medline: 27332249]
28. Partridge SR, Redfern J. Strategies to engage adolescents in digital health interventions for obesity prevention and management. Healthcare (Basel) 2018 Jun 21;6(3):70 [FREE Full text] [doi: 10.3390/healthcare6030070] [Medline: 29933550]
29. Chaplais E, Naughton G, Thivel D, Courteix D, Greene D. Smartphone interventions for weight treatment and behavioral change in pediatric obesity: a systematic review. Telemed J E Health 2015 Oct;21(10):822-830. [doi: 10.1089/tmj.2014.0197] [Medline: 26290954]
30. Golley R, Hendrie G, Slater A, Corsini N. Interventions that involve parents to improve children's weight-related nutrition intake and activity patterns - what nutrition and activity targets and behaviour change techniques are associated with intervention effectiveness? Obes Rev 2011 Feb;12(2):114-130. [doi: 10.1111/j.1467-789X.2010.00745.x] [Medline: 20406416]
31. Yavuz HM, van Ijzendoorn MH, Mesman J, van der Veek S. Interventions aimed at reducing obesity in early childhood: a meta-analysis of programs that involve parents. J Child Psychol Psychiatry 2015 Jun 08;56(6):677-692. [doi: 10.1111/jcpp.12330] [Medline: 25292319]
32. Young KM, Northern JJ, Lister KM, Drummond JA, O'Brien WH. A meta-analysis of family-behavioral weight-loss treatments for children. Clin Psychol Rev 2007 Mar;27(2):240-249. [doi: 10.1016/j.cpr.2006.08.003] [Medline: 17070638]
33. Niemeier BS, Hektner JM, Enger KB. Parent participation in weight-related health interventions for children and adolescents: a systematic review and meta-analysis. Prev Med 2012 Jul;55(1):3-13. [doi: 10.1016/j.ypmed.2012.04.021] [Medline: 22575353]

34. Nyström CD, Sandin S, Henriksson P, Henriksson H, Trolle-Lagerros Y, Larsson C, et al. Mobile-based intervention intended to stop obesity in preschool-aged children: the MINISTOP randomized controlled trial. Am J Clin Nutr 2017 Jun;105(6):1327-1335. [doi: 10.3945/ajcn.116.150995] [Medline: 28446496]
35. Leung MM, Mateo KF, Verdaguer S, Wyka K. Testing a web-based interactive comic tool to decrease obesity risk among minority preadolescents: protocol for a pilot randomized control trial. JMIR Res Protoc 2018 Nov 09;7(11):e10682 [FREE Full text] [doi: 10.2196/10682] [Medline: 30413399]
36. Dias JD, Domingues AN, Tibes CM, Zem-Mascarenhas SH, Fonseca LM. Serious games as an educational strategy to control childhood obesity: a systematic literature review. Rev Lat Am Enfermagem 2018 Sep 03;26:e3036 [FREE Full text] [doi: 10.1590/1518-8345.2509.3036] [Medline: 30183872]
37. Nguyen B, Kornman KP, Baur LA. A review of electronic interventions for prevention and treatment of overweight and obesity in young people. Obes Rev 2011 May;12(5):298-314. [doi: 10.1111/j.1467-789X.2010.00830.x] [Medline: 21348921]
38. Del Río NG, González-González CS, Martín-González R, Navarro-Adelantado V, Toledo-Delgado P, García-Peñalvo F. Effects of a gamified educational program in the nutrition of children with obesity. J Med Syst 2019 May 22;43(7):198. [doi: 10.1007/s10916-019-1293-6] [Medline: 31119385]
39. van der Aa DA, Altenburg TM, van Randeraad-van der Zee CH, Chinapaw MJ. The effectiveness and promising strategies of obesity prevention and treatment programmes among adolescents from disadvantaged backgrounds: a systematic review. Obes Rev 2017 May;18(5):581-593. [doi: 10.1111/obr.12519] [Medline: 28273680]

Abbreviations

mHealth: mobile health
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized clinical trial
SMD: standardized mean difference

Edited by R Kukafka; submitted 24.05.21; peer-reviewed by P Henriksson, J Alvarez Pitti, J Lee; comments to author 30.07.21; revised version received 10.09.21; accepted 22.09.21; published 14.02.22.

Please cite as:
Kouvari M, Karipidou M, Tsiampalis T, Mamalaki E, Poulimeneas D, Bathrellou E, Panagiotakos D, Yannakoulia M
Digital Health Interventions for Weight Management in Children and Adolescents: Systematic Review and Meta-analysis
J Med Internet Res 2022;24(2):e30675
URL: https://www.jmir.org/2022/2/e30675
doi:10.2196/30675
PMID:35156934

Review

Posthospitalization Follow-Up of Patients With Heart Failure Using eHealth Solutions: Restricted Systematic Review

Ingvild Margreta Morken[1,2*], PhD; Marianne Storm[3,4*], Prof Dr; Jon Arne Søreide[5,6*], Prof Dr; Kristin Hjorthaug Urstad[1,7*], Prof Dr; Bjørg Karlsen[3*], Prof Dr; Anne Marie Lunde Husebø[2,3*], Prof Dr

[1]Department of Quality and Health Technologies, University of Stavanger, Stavanger, Norway

[2]Research Group for Nursing and Health Sciences, Stavanger University Hospital, Stavanger, Norway

[3]Department of Public Health, University of Stavanger, Stavanger, Norway

[4]Faculty of Health Sciences and Social Care, Molde University College, Molde, Norway

[5]Department of Gastrointestinal Surgery, Stavanger University Hospital, Stavanger, Norway

[6]Department of Clinical Medicine, University of Bergen, Bergen, Norway

[7]Faculty of Health Studies, VID Specialized University, Oslo, Norway

[*]all authors contributed equally

Corresponding Author:
Anne Marie Lunde Husebø, Prof Dr
Department of Public Health
University of Stavanger
Prof Olav Hanssens vei 10
Stavanger, 4021
Norway
Phone: 47 99262805
Email: anne.m.husebo@uis.no

Abstract

Background: Heart failure (HF) is a clinical syndrome with high incidence rates, a substantial symptom and treatment burden, and a significant risk of readmission within 30 days after hospitalization. The COVID-19 pandemic has revealed the significance of using eHealth interventions to follow up on the care needs of patients with HF to support self-care, increase quality of life (QoL), and reduce readmission rates during the transition between hospital and home.

Objective: The aims of this review are to summarize research on the content and delivery modes of HF posthospitalization eHealth interventions, explore patient adherence to the interventions, and examine the effects on the patient outcomes of self-care, QoL, and readmissions.

Methods: A restricted systematic review study design was used. Literature searches and reviews followed the (PRISMA-S) Preferred Reporting Items for Systematic Reviews and Meta-Analyses literature search extension checklist, and the CINAHL, MEDLINE, Embase, and Cochrane Library databases were searched for studies published between 2015 and 2020. The review process involved 3 groups of researchers working in pairs. The Mixed Methods Appraisal Tool was used to assess the included studies' methodological quality. A thematic analysis method was used to analyze data extracted from the studies.

Results: A total of 18 studies were examined in this review. The studies were published between 2015 and 2019, with 56% (10/18) of them published in the United States. Of the 18 studies, 16 (89%) were randomized controlled trials, and 14 (78%) recruited patients upon hospital discharge to eHealth interventions lasting from 14 days to 12 months. The studies involved structured telephone calls, interactive voice response, and telemonitoring and included elements of patient education, counseling, social and emotional support, and self-monitoring of symptoms and vital signs. Of the 18 studies, 11 (61%) provided information on patient adherence, and the adherence levels were 72%-99%. When used for posthospitalization follow-up of patients with HF, eHealth interventions can positively affect QoL, whereas its impact is less evident for self-care and readmissions.

Conclusions: This review suggests that patients with HF should receive prompt follow-up after hospitalization and eHealth interventions have the potential to improve these patients' QoL. Patient adherence in eHealth follow-up trials shows promise for successful future interventions and adherence research. Further studies are warranted to examine the effects of eHealth interventions on self-care and readmissions among patients with HF.

(J Med Internet Res 2022;24(2):e32946) doi:10.2196/32946

KEYWORDS

adherence; eHealth; heart failure; posthospitalization follow-up; patient outcome; review

Introduction

Background

Heart failure (HF) affects an estimated 64 million people worldwide [1]. It poses a burden on the health care system in general and on primary care specifically because the total number of patients with HF is increasing, reflecting the chronic course of the disease as well as population growth and aging [2,3]. Symptomatic HF is a complex clinical syndrome with a symptom burden of dyspnea and fatigue [4] and can be troublesome for patients and their families because of frequent hospitalizations and symptom and treatment burden negatively affecting their quality of life (QoL) [5-7]. QoL is understood as a multidimensional and subjective concept that includes physical, functional, emotional, and social well-being [8]. Effective self-care behavior is essential for patients with HF [9,10]. Self-care in the context of HF is an overarching concept based on three key concepts: (1) self-care maintenance (eg, compliance with medication regimens and following diet and physical activity recommendations), (2) self-care monitoring (eg, regular weighing), and (3) self-care management (eg, changing diuretic dose in response to symptoms) [10]. Upon discharge from the hospital, many patients transition from care provided by health professionals in a safe hospital setting to individual self-care at home [11]. This period, when patients transition between hospital and home, is a vulnerable and stressful time for patients with HF and many struggle to perform recommended self-care and navigate the health care system, particularly when posthospitalization care is poorly executed as a result of inadequate coordination of resources or follow-up [5,7]. Of any diagnosis, HF is associated with the highest 30-day all-cause readmission rate (approximately 20%), whereas nearly 35% of the patients with HF are readmitted within 90 days [2,6]. During this phase, the lack of resources for following up or poor medical education leaves this population vulnerable to deterioration and rehospitalization [12]. Posthospitalization HF disease management programs include education, self-management, weight monitoring, sodium restriction or dietary advice, exercise recommendations, and medication review [13]. In addition to social and psychological support with a high degree of care coordination, as well as the higher intensity of follow-up, these components may be important for better self-care behavior, increased QoL, and reduced readmission rates [4,13,14]. The impact of the COVID-19 pandemic has raised the requirement for, and importance of, eHealth solutions as a tool for health care professionals to perform such follow-up of patients with HF [15]. Insight into ensuring a more seamless eHealth care service from inpatient to outpatient care for patients with HF is necessary if they are to achieve adequate self-care support and feel safe [15,16]. eHealth care service is defined as "health services and information delivered or enhanced through the internet and related technologies" [17] and holds the potential to increase the efficiency and quality of health care services [18]. In this review, eHealth comprises digital solutions to deliver health care services, including patient education; telemonitoring of weight, blood pressure, and heart rhythm; and social and emotional support. Previous research suggests that the use of posthospitalization eHealth interventions to follow up on patients may promote self-care for people with long-term illness [18]. Several recent reviews have summarized the findings from eHealth follow-up interventions for patients with HF and provided information about the efficiency of such interventions. Auener et al [19] investigated the effects of telemonitoring programs on different aspects of health care use from 16 randomized controlled trials (RCTs) and 13 nonrandomized studies. All studies included weight as a parameter, whereas only 4 included electrocardiography measures as a physiological parameter. The results revealed that telemonitoring has the potential to reduce hospitalization rates. However, the number of non–emergency department visits increased in most of the studies [18]. Ding et al [20] extracted 18 telemonitoring strategies from 26 RCTs involving patients with HF. Some strategies were commonly used, such as call center support and daily weight monitoring, whereas others, including nurse support, interventions for depression and anxiety, and exercise interventions, were seldom used. Telemonitoring strategies involving medication support and mobile health (mHealth) interventions were associated with improvements in all-cause mortality or hospitalization outcomes [20]. A systematic review conducted in 2017 identified 39 relevant RCTs of telemedicine, largely based on assessments of symptoms, weight, heart rate and rhythm, and blood pressure, and found that telemonitoring was associated with reductions in all-cause mortality of 20% and HF hospitalization of 37% [21]. In contrast, nurse-based telephone-supported care seemed to provide little benefit, and only a reduction in the rate of HF-related admission was noted compared with the control group. However, a combination of home-based teletransmission and nurse-based telephone reinforcement may be encouraged [21]. Although these reviews generally support the effectiveness of eHealth interventions for patients with HF, the outcomes mainly focus on readmission and health care use, and only one of them focuses specifically on the hospital-to-home transition phase [21]. Moreover, they mostly lack information about self-management, QoL, and participants' adherence to the eHealth interventions. Adherence to self-management and medication regimens is crucial during the transition from hospital discharge to home to prevent hospital readmission and achieve improved health outcomes and QoL [22-24]. Therefore, the success of an intervention aiming to support patients' chronic disease management depends on patient adherence to the intervention components [25]. Intervention adherence refers to the degree to which the behavior of trial participants corresponds to the intervention assigned to them [26]. Adherence varies according to the patient's health status, treatment regimens, access to support, and psychological factors such as motivation and beliefs. The long-term success of interventions depends on patients assuming responsibility for their own health and can

be achieved with the aid of coordinated measures such as patient education and regular follow-up contacts [26]. An accurate assessment of intervention adherence is warranted to verify whether changes in health outcomes are due to a particular intervention [26].

There is a knowledge gap concerning the synthesis of recent posthospitalization eHealth follow-up interventions for patients with HF focusing on outcomes of self-care, QoL, and adherence to the interventions. Therefore, this restricted review will investigate eHealth interventions that may better prepare patients for the period after hospital discharge, strengthen their self-care and QoL, reduce readmissions, and help them to recover well. Furthermore, the review will address the issue of adherence and discuss how it may affect intervention outcomes. Therefore, the aim is to summarize the most recent information about the content and delivery mode of HF posthospitalization eHealth interventions, explore patient adherence to the interventions, and systematically investigate the effects on patient outcomes of self-care, QoL, and readmissions.

Research Questions

The a priori research questions were designed according to the FINER framework, which states that a review research question should be feasible, interesting, novel, ethical, and relevant [27].

Our research questions were as follows:

- What are the content and delivery modes of posthospitalization eHealth interventions for patients with HF?
- What is the reported adherence to posthospitalization eHealth interventions in HF?
- Which effects can be expected from posthospitalization eHealth interventions on self-care, QoL, and readmissions of patients who have received treatment for HF?

Methods

Reporting Standards

This study used a framework proposed for restricted systematic reviews [27]. The restricted systematic review framework is proposed to be applicable when conducting a rapid review because it consists of core steps that are minimum requirements for systematic reviews, thereby accommodating factors such as a short time frame and limited resources [28]. Such factors are important to consider when conducting a literature search and review as part of developing complex interventions [29]. The framework comprises six core steps: (1) literature search, (2) study selection, (3) data extraction, (4) critical assessment of the included studies, (5) data synthesis, and (6) publication [28].

Step 1: Literature Search and Search Terms

The literature search was performed as part of a more extensive review study on eHealth interventions in noncommunicable diseases. This paper reports the results from HF populations. A research librarian performed comprehensive literature searches in the CINAHL, MEDLINE, Embase, and Cochrane Library databases. To ensure that our results reflect current conditions and avoid repeating previous review efforts, this rapid review was limited to data published between 2015 and 2020 in English or a Scandinavian language. Searches were performed in the publication title or abstract. Appropriate search terms, including relevant Medical Subject Headings, were closely matched with the Population, Intervention, Control, and Outcome elements (see next section). Documentation of the search strategy and search terms is presented in Multimedia Appendix 1. The search strategy also included manually hand searching the reference lists of the included studies and relevant background material. The searches were performed on March 20, 2020.

A Priori Eligibility Criteria

Key components of the synthesis are encapsulated by the Population, Intervention, Control, and Outcome framework [30].

- Population: patients initially treated for HF
- Intervention: posthospitalization eHealth follow-up services
- Control: standard care and nondigital follow-up services
- Outcomes: self-management and self-care, QoL, and readmissions

The inclusion and exclusion criteria are displayed in Textbox 1.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Empirical intervention studies
- Populations of adult patients with heart failure
- eHealth interventions from hospital to home
- Patient outcomes of self-care, quality of life, and readmissions
- Experimental and quasi-experimental randomized and nonrandomized controlled trials
- Pre–post design with a comparison group
- Peer-reviewed studies
- Published in English

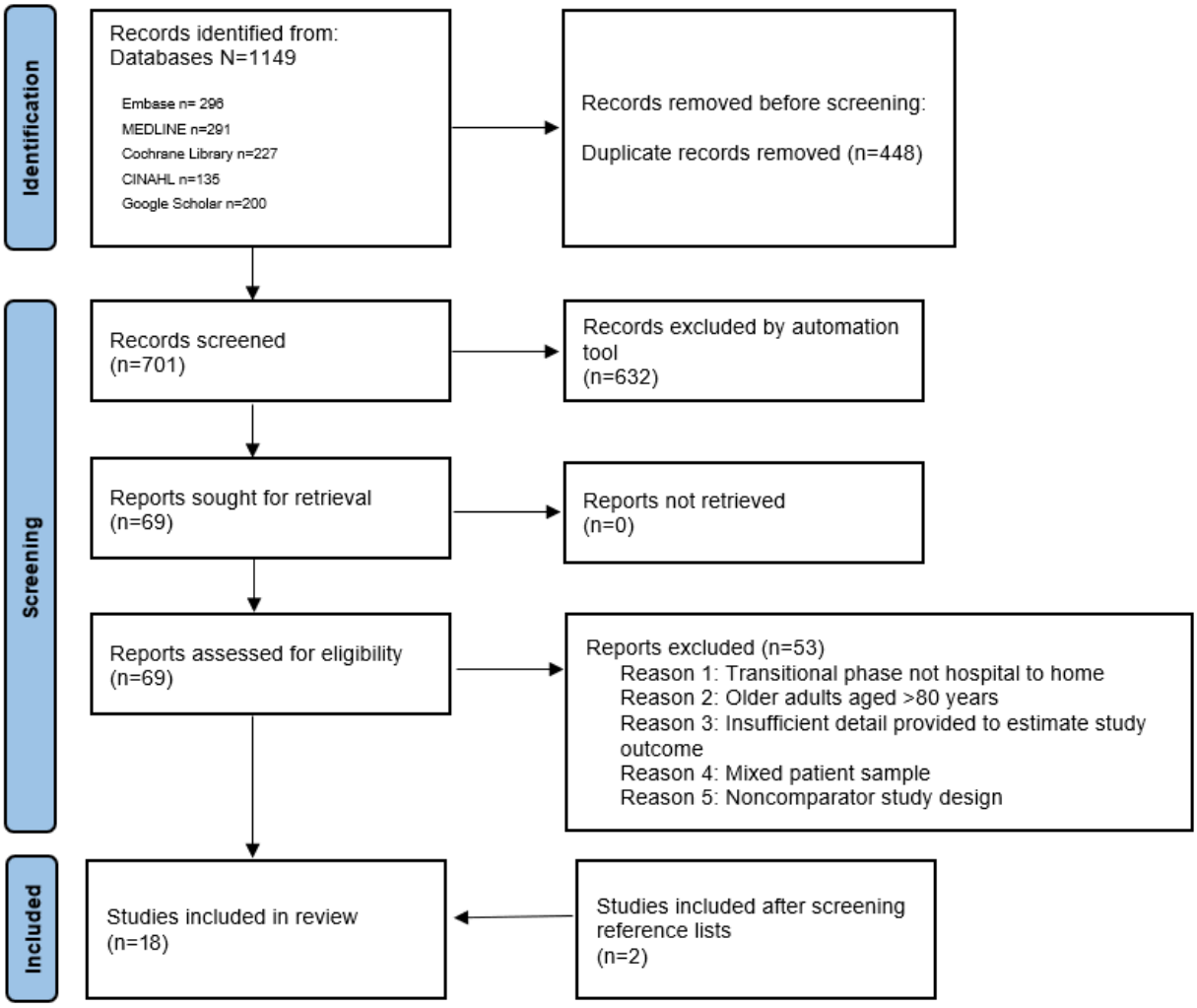
Exclusion criteria

- Review studies, study protocols, book chapters, and conference contributions
- Children and adolescent patients
- Older adults (aged >80 years)
- Community health care services context
- >3 months since hospital discharge
- Insufficient detail provided to estimate study outcome
- Mixed patient samples
- Noncomparator study designs

Step 2: Study Selection

After removing duplicates using EndNote (Clarivate), a member of the research team (AMLH) carried out an initial broad review of all included titles and abstracts, using the a priori inclusion and exclusion criteria. Next, the abstracts verified for potential inclusion were reviewed for full-text extraction by all authors, divided into 3 review teams. Full-text articles were extracted for 9.8% (69/701) of the abstracts. Finally, team members resolved conflicting opinions by assessing reasons for exclusion and deciding whether to include the study. The results of the data search and selection process are displayed in a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart (Figure 1) [31].

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow chart of the study selection process.



Step 3: Data Extraction

An Excel spreadsheet (Microsoft Corp) was created to ensure consistent data extraction, including data fields of publication identifiers, study design, study context and participants, eHealth intervention or program, and outcomes. The review teams used the spreadsheet to extract relevant data from the included articles. Any inconsistency within the group was resolved through assessment by a reviewer from one of the other groups.

Step 4: Critical Assessment of Included Studies

To minimize bias, an assessment of internal validity of the included studies, risk of bias (eg, over- or underestimation of intervention effect), and potential conflicts of interest were examined using the Mixed Methods Appraisal Tool (MMAT) [32]. The MMAT, which aims to appraise the methodological quality of included studies in systematic reviews, consists of a checklist of qualitative, quantitative, and mixed methods studies [32]. For this review, checklists for randomized and nonrandomized research designs were used. Each checklist is initiated with 2 screening questions to allow for further assessment, and each list contains 5 assessment criteria to be answered with *Yes*, *No*, or *Can't tell*. A total score of 7 constitutes a *Yes* response to the screening and assessment criteria [32]. The developers recommend that the MMAT be used to describe only the study quality and to avoid excluding studies based on total scores [32].

To assess data quality, each review team member independently rated the studies, followed by a discussion to achieve consensus. For 10% (2/18) of the included studies, the quality scoring was verified through independent scoring by 2 reviewers (IMM and AMLH). The quality of included studies was above moderate (ie, of the 7 criteria, 6 [86%] were answered with *Yes*; Textbox 1).

Step 5: Data Synthesis

The findings on service content and delivery mode, adherence, and the effects of posthospitalization follow-up eHealth interventions were systematically analyzed by using thematic analysis as well as searching for patterns, themes, and categories across studies, which were then narratively summarized as suggested by Whittemore and Knafl [33]. Because of the heterogeneity of the study designs, participants, and outcome measures, meta-analysis was not recommended. Thus, the effects on patient outcomes were reviewed and reported narratively.

Step 6: Publication

The results from this restrictive systematic review will be published, including all appendices and added data. In addition, the study's findings will be disseminated in relevant clinical settings and websites.

Results

Overview

The literature search process is outlined in Figure 1. The search yielded a total of 1149 references (ie, records screened for 2 patient populations); after the removal of 318 (27.68%) duplicates, 831 (72.32%) titles and abstracts were assessed for inclusion. Of the 831 titles and abstracts, 701 (84.4%) titles pertaining to eHealth interventions for patients with HF were screened for eligibility using the web tool [34]. Of the 69 studies evaluated for eligibility in full text, 16 (23%) met all inclusion criteria and were included. Screening the reference lists of the included studies yielded another study and screening the reference lists of relevant background material identified a further study. Finally, this review included 18 studies.

Study Characteristics

Detailed characteristics of the included studies are displayed in Table 1. All studies were published between 2015 and 2019. Of the 18 studies, 10 (56%) were performed in the United States [35-44]. Although RCT was the predominant study design, 11% (2/18) of the studies applied a quasi-experimental method [37,43]. Among the 18 studies, enrollment of patients with HF to the posthospitalization eHealth service varied from recruitment at the hospital before hospital discharge to recruitment within 3 months of recent hospitalization (Textbox 1). In 56% (10/18) of the studies, all patients were recruited upon hospital discharge to an intervention with a duration of 14 days to 12 months [35,37,38,40-42,44-47]. In 22% (4/18) of the studies, patients were recruited after recent (within 30 days) hospitalization to an intervention with a 3- to 12-month duration [39,43,48,49]. In another study, patients with HF were enrolled during hospitalization or within 3 months of discharge for an HF exacerbation, and the intervention lasted for 3 months [36]. In 17% (3/18) of the studies, patients were recruited at hospital discharge or at the HF outpatient clinic to an intervention with a duration of 3-9 months [50-52].

Table 1. Characteristics of included eHealth intervention studies involving patients with heart failure (HF; N=18).

Study (country)	Design	Sample size			Content, focus, and mode of instruction	Duration	MMAT[a] scores out of 7, n (%)
		Total sample (N)	I[b], n (%)	C[c], n (%)			
Athilingam et al [35] (United States)	RCT[d]	18	9 (50)	9 (50)	Telemonitoring (HeartMapp); daily measures of weight, heart rate, blood pressure, and HF symptoms. HF education: 10 modules, home visit after 2-3 days by a nurse. A phone call to all participants. Nurses checked the dashboard daily to monitor participants' progress.	30 days	2 (29)
Comin-Colet et al [45] (Spain)	RCT	178	81 (45.5)	97 (54.5)	Telemonitoring and telephone support. Daily measures of weight, heart rate, and blood pressure. HF nurses reviewed alarms and alerts from the system every day.	6 months	6 (86)
Dunbar [36] (United States)	RCT	134	70 (52.2)	64 (47.8)	Telephone support; education and counseling on diet, medications, self-monitoring, symptoms, and physical activity; self-monitored blood glucose level and weight; self-care with follow-up home visits and telephone counseling.	6 months	4 (57)
Evangelista et al [37] (United States)	Quasi-experimental	42	21 (50)	21 (50)	Telemonitoring and telephone support; daily measures of weight, heart rate, and blood pressure. Telemonitoring provided alerts and feedback in the case of worrisome responses to questions or if vital signs were outside of preset limits. The research nurse communicated with the patient through teleconferencing and collaborated with the patient's primary care provider to facilitate a plan of action. Telephone support as usual to the control group.	3 months	7 (100)
Frederix et al [46] (Belgium)	RCT	160	80 (50)	80 (50)	Telemonitoring; daily measurements of weight, heart rate, and blood pressure were forwarded to a central computer. If the recordings were outside of predefined alert limits, both the general practitioner and HF clinic were alerted by email. At that moment, per protocol, the general practitioner (or cardiologist) was asked to visit or contact the patient and adapt the treatment if they felt that it was necessary. The HF nurse contacted the patient by telephone 1-3 days after the alert to verify whether the intervention had been effective.	6 months	6 (86)
Gallagher et al [38] (United States)	RCT	40	20 (50)	20 (50)	Telemonitoring; electronic measurement of adherence to loop diuretics. A licensed clinical social worker reviewed adherence data daily during the first 7 days after discharge and weekly after that and then contacted participants who were nonadherent for ≥2 days per week.	30 days	7 (100)

Study (country)	Design	Sample size			Content, focus, and mode of instruction	Duration	MMAT[a] scores out of 7, n (%)
		Total sample (N)	I[b], n (%)	C[c], n (%)			
Hwang et al [48] (Australia)	RCT	53	24 (45.3)	29 (54.7)	Telemonitoring and telephone support; participants were instructed to self-monitor and verbally report their blood pressure, heart rate, and oxygen saturation levels at the start of each rehabilitation session. The intervention group received electronic education sessions.	3 months	7 (100)
Jayaram et al [39] (United States)	RCT	1521	756 (49.7)	765 (50.3)	Telephone calls are used for technical support by interactive voice response; symptoms and daily weight; patients were instructed to call a toll-free number daily for 6 months, respond to a series of automated questions regarding their symptoms, and enter their daily weight. They were also provided with educational materials.	6 months	6 (86)
Kotooka et al [47] (Japan)	RCT	181	90 (50)	91 (50)	Telemonitoring and telephone support, measurement of weight, heart rate, and blood pressure daily. Physicians could provide telephone guidance, change medications, and order hospital readmission if required. Full-time nurses monitored acquired data on a secure website. Telephone support from a physician as usual.	15 months	6 (86)
Kraai et al [50] (Netherlands)	RCT	176	83 (47.2)	93 (52.8)	Telemonitoring and telephone support; daily measurement of weight, heart rate, and blood pressure. HF nurses automatically received notifications by mobile phone and email and then discussed symptoms and treatment with patients within 2 hours.	9 months	6 (86)
Köberich et al [51] (Germany)	RCT	110	58 (52.7)	52 (47.3)	Telephone support; nurse-led symptom monitoring, education on signs and symptoms of worsening HF, HF-specific diet, and fluid restriction. When seeking help, patients were advised to use a diary to document body weight, blood pressure, heart rate, and edema on a daily basis. If necessary, after discharge from the hospital, patients received 4 telephone calls within 3 months about changes in HF-related symptoms and treatment.	3 months	5 (71)
Lycholip et al [49] (Netherlands)	RCT	118	58 (49.2)	60 (50.8)	Telemonitoring and telephone support; daily measurement of body weight, blood pressure, and heart rate. HF nurses automatically received notifications by mobile phone and email and, within 2 hours, discussed the symptoms and treatment with the patients. An HF nurse provided education on HF.	9 months	6 (86)

Study (country)	Design	Sample size			Content, focus, and mode of instruction	Duration	MMAT[a] scores out of 7, n (%)
		Total sample (N)	I[b], n (%)	C[c], n (%)			
Masterson- Creber et al [40] (United States)	RCT	67	41 (61.2)	26 (38.8)	Telephone support MI[e]: a tailored intervention at discharge to improve self-care, involving a home visit and follow-up calls. A nurse used the MI approach to identify client-directed self-care goals. Participants received written educational material.	3 months	6 (86)
Ong et al [41] (United States)	RCT	1437	715 (49.7)	722 (50.3)	Telemonitoring and telephone support; weight, heart rate, and blood pressure were measured daily. A total of 9 telephone health coaching calls over 6 months, generally from the same call center nurse.	6 months	5 (71)
Pedone et al [52] (Italy)	RCT	96	50 (52/1)	46 (47.9)	Telemonitoring and telephone support; measurement of blood pressure, oxygen saturation, weight, and heart rate daily; a geriatrician evaluated the data received every day. Participants received education on medical treatment and lifestyle counseling by telephone.	6 months	6 (86)
Ritchie et al [42] (United States)	RCT	511	253 (49.5)	258 (50.5)	Interactive voice response and telephone support; symptoms and body weight measured daily; E-Coach intervention: an intervention with condition-specific customization and in-hospital and postdischarge support by a care transition nurse, interactive voice response, postdischarge calls, and care transition nurse follow-up.	2 months	7 (100)
Srivastava et al [43] (United States)	Cohort–control	1067	197 (18.5)	870 (81.5)	Telemonitoring and telephone support; measurement of heart rate and blood pressure daily. Data were monitored on weekdays by a telehealth nurse who analyzed the data for abnormalities and lack of response; if clinical data caused concern for declining health status, a phone call was initiated to the patient. All patients also received a monthly follow-up call.	12 months	6 (86)
Young et al [44] (United States)	RCT	105	54 (51.4)	51 (48.6)	Telephone support: the patient-activated care at home intervention contained a variety of formats (eg, verbal, written, and visual) with 12 weeks of post discharge education sessions delivered by telephone. Besides self-management workbooks, each subject was provided with a self-management toolkit, including a calendar for weight and daily salt-intake logging, a step-on weight scale with large and bright readings, and an electronic pill organizer reminder alarm.	6 months	6 (86)

[a]MMAT: Mixed Methods Appraisal Tool.

[b]I: intervention.

[c]C: control.

[d]RCT: randomized controlled trial.

[e]MI: motivational interview.

Themes Derived From Data Analysis

In the following section, the data analysis results are presented, thereby answering the research questions concerning intervention content and delivery mode, intervention adherence, and the effects of eHealth on patient outcomes.

Delivery Mode and Content of Posthospitalization eHealth Follow-Up Interventions

In all, 2 different modes of delivering an eHealth service were identified (Table 1). The specific technologies identified included (1) structured telephone calls and (2) telemonitoring or telemonitoring in combination with telephone support.

Structured Telephone Call

Of the 18 studies in our review, 6 (33%) included structured telephone calls to deliver the intervention to patients with HF [36,39,40,42,44,51]. Of these 6 studies, in 2 (33%), interactive voice response devices were used to examine the patients' symptoms and vital sign registrations [39,42]. In these studies, patients were instructed to call a toll-free number daily for 6 months, respond to a series of automated questions about their symptoms, and enter their daily weight. Responses that met prespecified criteria triggered a variance within the system. Conflicts were then flagged for immediate attention by on-site clinicians [39].

Nurses performed all telephone calls. Four dominant categories of content and use of the telephone-supported HF interventions were identified as follows: (1) keeping logs: encouraging patients to keep logs for monitoring symptoms, blood pressure, and weight; (2) goal-setting skills: teaching patients goal-setting skills to manage their condition or behavior changes; (3) problem-solving skills: teaching patients problem-solving skills to manage their condition; and (4) advice about when to seek help in case of worsening HF. In addition, education and counseling were combined with follow-up home visits in 17% (1/6) of the studies [36], whereas in another study, customized HF education was provided on the patient's response to questions on symptoms and self-management [40]. Each intervention session lasted 15-50 minutes. In the trial conducted by Ritchie et al [42], support calls were provided to patients only when required, whereas in 67% (4/6) of the studies, 4-10 calls were delivered for 2-4 months [36,40,44,51].

Telemonitoring

Of the 18 included studies, 12 (67%) included a telemonitoring program. In 75% (9/12) of these studies, weight, heart rate, and blood pressure were measured daily [35,37,41,45-47,49,50,52]. Athilingam et al [35] also included a medication tracker in their HeartMapp app and physiological exercises to reset the autonomic nervous system and improve functional capacity. Pedone et al [52] included oxygen saturation in addition to measuring blood pressure and heart rate daily. These studies also used assessments of symptoms related to HF and action plans for clinical decisions based on out-of-limit alerts from the data monitoring. In 75% (9/12) of the studies, nurses specialized in HF care and telemedicine, or care transition performed the daily data monitoring [35,37,41,43,45-47,49,50]. Of the 12 telemonitoring studies, 2 (17%) provided patients with automated feedback triggered by out-of-limit alerts [35,39]. In cases where these alerts indicated possible mild to moderate decompensation, nurses could promote diuretic dose adjustments following specific protocols [45] and alerts could be routed to clinicians (eg, physicians and cardiologists) who evaluated the data and contacted patients if necessary. For cases in which out-of-limit alerts indicated severe decompensation, patients were advised to call the emergency number or go to the nearest hospital emergency department [35,37,41,43,45-48,50]. In 17% (2/12) of the studies, clinicians (physicians and geriatricians) conducted data monitoring and management simultaneously [39,52]. Of the 12 studies, 1 (8%) was a telerehabilitation investigation in which participants were guided to self-monitor and verbally report their blood pressure, heart rate, and oxygen saturation levels at the start of each rehabilitation session [48]. Finally, 75% (9/12) of the telemonitoring studies provided the participants with telephone support to either follow up on alerts generated from the patient's registrations of symptoms and vital signs [43,45,49,50], technical support [48], and follow-up of control group or as usual care [37,47] or to provide patient education [41,52].

Adherence to Posthospitalization Follow-Up eHealth Interventions in HF

Of the 18 included studies, 11 (61%) reported patients' adherence to the intervention [35,38,41,42,44,45,47-50,52]. In 91% (10/11) of these studies, adherence was reported as a secondary study outcome, whereas in 9% (1/11) of the studies, adherence was included as a primary outcome [38]. Overall, adherence levels were reported at a rate of 72%-99%. Among the studies using telephone support as a delivery mode, 33% (2/6) included measures of adherence with adherence levels of 86% [42] and 84% (T1) and 86% (T2) [44]. Details regarding reported adherence are provided in Table 2.

Table 2. Reporting intervention program adherence in the included studies (N=18).

Study	Adherence reported	Definition and assessment of adherence	Adherence results
Athilingam et al [35]	Yes	Duration for which the participants accessed intervention features.	Adherence was low, with only 72% of the participants completing the 30-day follow-up.
Comin-Colet et al [45]	Yes	Daily automated telemonitoring of biometrics and symptoms using the intervention platform.	Adherence was very high, with missed biometric daily transmissions less than 1% of the expected number of daily transmissions.
Dunbar et al [36]	No	—[a]	—
Evangelista et al [37]	No	—	—
Frederix et al [46]	No	—	—
Gallagher et al [38]	Yes	Adherence to loop diuretics in the 30 days after discharge. Nonadherence=adherence <88%. Adherence was calculated as the percentage of days on which the correct number of doses was taken as prescribed, irrespective of dose timing.	Median correct dosing adherence was 81%, and 33% of the participants were classified as adherent. Reasons for nonadherence were identified as follows: ran out of pills, out of usual routine, side effects, and did not know the correct dose.
Hwang et al [48]	Yes	Attendance rates were categorized into adherent (>80%), partly adherent (20%-80%), and nonadherent (<20%), based on the proportion of sessions attended by each participant.	Of the 51 participants who attended the rehabilitation programs, 49 (96%) were categorized as adherent or partly adherent. None of the intervention participants were nonadherent.
Jayaram et al [39]	No	—	—
Köberich et al [51]	No	—	—
Kotooka et al [47]	Yes	Adherence was measured as the number of days that each patient measured their body weight and blood pressure in a month.	The mean rates of adherence at 1, 6, and 12 months after randomization were 96%, 90%, and 91%, respectively.
Kraai et al [50]	Yes	Adherence of patients to telemonitoring was assessed by daily weighing and measuring of blood pressure.	The median adherence rate was 95% (range 87%-99% for the total study period).
Lycholip et al [49][b]	Yes	Adherence of patients to telemonitoring was assessed by daily weighing and measuring of blood pressure.	The median adherence rate was 95% (range 87%-99% for the total study period).
Masterson-Creber et al [40]	No	—	—
Ong et al [41]	Yes	Telemonitoring adherence: percentage of total days during 30 and 180 days; telephone coaching adherence: percentage of total days during 30 and 180 days.	Overall, 83% (591/715) of the intervention participants used telemonitoring equipment.
Pedone et al [52]	Yes	Percentage of the total amount of expected symptom measurements.	On average, 62% of the scheduled measurements were completed (weight once a day, blood pressure and heart rate twice a day, and peripheral oxygen saturation thrice a day); adherence was best for pulse oximeter (70%) and worst for the scale (56%); 64% of the participants completed at least half of the scheduled measurements.
Ritchie et al [42]	Yes	Total (100%) adherence was defined as answering all interactive voice response system calls. Optimal adherence: daily response to the interactive voice response during the first 7 days. Answering a call was defined as a patient completing the questions of the call.	Of the patients with HF, 144 (86%) received a total intervention dose.
Srivastava et al [43]	No	—	—
Young et al [44]	Yes	Frequencies of self-reported self-management behaviors of daily weighing, following a low-sodium diet, taking prescribed medications, exercising, and attending follow-up appointments.	Participants in the intervention group who received the patient-activated care at home intervention had significantly higher self-reported adherence to self-management behaviors; 84% at 3 months and 86% at 6 months reported not missing any doses in the previous week, compared with 68% at both time points in the control group.

[a]Data not available.

[b]Same study population and intervention as in the study by Kraai [50].

Effects From Follow-Up Interventions on Patient Outcomes

Overview

Of the 18 included studies, only 1 (6%) investigated all 3 patient outcomes of interest to this review (ie, QoL, readmissions, and self-care behavior) [45]. Included in 61% (11/18) of the studies, QoL was the most frequently analyzed patient outcome, followed by readmissions in 56% (10/18) of the studies. Self-care was explored in 44% (8/18) of the included studies. Details concerning the effects of eHealth interventions are provided in Table 3.

Table 3. Effects of intervention programs on patient outcomes of quality of life (QoL), self-care, and readmissions (N=18).

Study	Sample size n (%), I[a]/C[b]	Baseline	Postbaseline measures		Outcome
			T1[c] (days), P value	T2[d] (days), P value	
Athilingam et al [35]	9/9 (50/50)	Hospital discharge	• .93 (30) • .01 (30) • .03 (30) • .18 (30)	N/A[e]	• Self-care maintenance • Self-care management • Self-care confidence • QoL
Comin-Colet et al [45]	81/97 (46/54)	Hospital discharge	• .06 (180) • .001 (180) • .01 (180)	N/A	• Self-care • QoL • Readmissions
Dunbar et al [36]	54/54 (50/50)	Hospital discharge or within 3 months after discharge	• <.001 (90)	• .002 (180)	• QoL
Evangelista et al [37]	21/21 (50/50)	Hospital discharge	• <.001 (90) • <.001 (90) • <.001 (90) • <.001 (90) • <.001 (90)	N/A	• QoL overall • QoL emotional subscale • Self-care maintenance • Self-care management • Self-care confidence
Frederix et al [46]	80/80 (50/50)	Hospital discharge	• .04 (180)	N/A	• Readmissions
Gallagher et al [38]	20/20 (50/50)	Hospital discharge	• .41 (30) • .72 (30)	N/A	• Self-care (medication adherence) • Readmissions
Hwang et al [48]	24/26 (48/52)	Recent discharge	• .03 (360)	• .03 (720)	• QoL
Jayaram et al [39]	756/765 (49.7/50.3)	Recent discharge	• .32 (90)	• .04 (180)	• QoL
Kotooka et al [47]	90/91- (50/50)	Hospital discharge	• .94 (352) • .42 (352)	N/A	• QoL • HFf readmissions
Kraai et al [50]	94/83 (53/47)	Hospital discharge or outpatient clinic	• .62 (270) • .87 (270)	N/A	• QoL • HF readmissions
Köberich et al [51]	58/52 (53/47)	Hospital discharge or outpatient clinic	• .20 (90) • <.001 (90)	N/A	• QoL • Self-care
Lycholip et al [49]	58/60 (49/51)	Recent discharge	• .77 (90)	N/A	• Self-care
Masterson-Creber et al [40]	41/26 (61/39)	Hospital discharge	• .36 (90) • .03 (90) • .31 (90)	N/A	• QoL • Self-care maintenance • Self-care confidence
Ong et al [41]	715/722 (49.8/50.2)	Hospital discharge	• .25 (30) • .63 (30)	• .02 (180) • .54 (180)	• QoL • Readmissions
Pedone et al [52]	50/46 (52/48)	Hospital discharge or outpatient clinic	• .04 (180)	N/A	• Readmissions
Ritchie et al [42]	245/233 (51.3/48.7)	Hospital discharge	• .18 (30)	N/A	• Readmissions
Srivastava et al [43]	197/870 (18.5/81.5)	Recent discharge	• .07 (352)	N/A	• Readmissions
Young et al [44]	54/51 (51/49)	Hospital discharge	• .09 (90) • <.001 (90)	• .09 (180) • <.001 (180)	• Readmissions • Self-care adherence

[a]I: intervention.

[b]C: control.

[c]T1: first postbaseline data collection.

[d]T2: second postbaseline data collection.

[e]N/A: not applicable.

[f]HF: heart failure.

Impact on QoL

QoL was included as a patient outcome in 61% (11/18) of the studies among patients with HF [35-37,39-41,45,47,48,50,51]. Of these 11 studies, 4 (36%) found that an eHealth intervention significantly improved patients' *overall QoL* [36,37,45,48]. Of these 4 studies, 3 (75%) contained telemonitoring combined with telephone support and with an intervention duration of 3-6 months [35,43,45], whereas the study by Dunbar et al [36] provided only telephone support lasting for 6 months. In both the studies by Jayaram et al [39] and Ong et al [41], the QoL was nonsignificant at the first postbaseline data collection. In contrast, QoL was significantly improved in both studies 6 months after beginning the intervention [39,41]. Of the 3 studies recruiting participants later than at discharge (ie, within 30 days after discharge), 2 (67%) reported significant effects from an eHealth intervention on QoL [39,48].

Self-care Behavior

Self-care was investigated in 39% (7/18) of the studies [35,37,38,44,45,49,51]. Athilingam et al [35] and Evangelista et al [37] both reported on self-care measured by the Self-Care of Heart Failure Index, which included the subscales self-care maintenance, self-care management, and self-care confidence. *Self-care management* was found to be significantly increased by eHealth interventions in both studies; in addition, Evangelista et al [37] found that *self-management maintenance* also seemed to be significantly improved. The subscale *self-care confidence* was enhanced to a significant degree by an eHealth intervention in 14% (1/7) of the studies [35]. Comin-Colet et al [45] used the Self-care Behavior Scale to study self-care behavior in patients with HF who were remotely followed by the Home Tele-HealthCare platform, with the authors detecting a marginally significant difference between the intervention and control groups. Köberich et al [51] and Lycholip et al [49] measured self-care behavior by using the European Heart Failure Self-care Behavior Scale. Of the 7 studies, 1 (14%) found significant improvements from the eHealth interventions on self-care behavior [51], whereas Lycholip et al [49] determined that such interventions did not influence self-care behavior; this study recruited patients within 14 days after discharge [49]. Finally, in the study by Gallagher et al [38], self-care was defined as medication adherence, with no significant effect from the eHealth intervention being noted. The studies showing significant effects on self-care behavior delivered the interventions for 30 days to 6 months. Of the 7 studies, only 1 (14%) did not include digital monitoring of symptoms and vital signs [46].

Readmissions

Of the 18 studies, 10 (56%) included readmissions as a patient outcome [38,41-47,50,52]. A significant reduction in readmissions associated with eHealth interventions was detected in 30% (3/10) of these studies [45,46,52]. All the studies combined telemonitoring and telephone support as the intervention delivery mode, and the intervention lasted for 6 months. Comin-Colet et al [45] found a significant reduction in readmissions in the HF intervention group compared with controls. The study by Frederix et al [46] identified a significant reduction (P=.04) in days lost to HF-related readmissions among patients in the intervention group but not for all-cause readmissions (P=.26). Pedone et al [52] revealed a significantly (P=.04) higher risk of readmissions (42%) at 180 days in the control group compared with 21% for patients with HF who were given remote follow-up. None of the studies that recruited patients later than discharge achieved significant effects on readmissions [43,44,47].

Discussion

Summary of Evidence and Comparison With Prior Work

In this restricted systematic review, we have evaluated and synthesized the findings from 18 posthospitalization follow-up eHealth interventions targeting QoL outcomes, self-care, and readmissions of patients with HF. To summarize, patients with HF were enrolled in the interventions upon or after hospital discharge. Interventions were delivered mainly by telephone or email and focused on patient education and counseling, social and emotional support, and self-monitoring of vital signs and symptoms. Posthospitalization eHealth follow-up for patients with HF holds potential for improving their QoL, whereas a positive impact on self-care and readmissions is less evident.

Some of the included studies used more traditional tools to follow up with patients, such as the telephone. Because of its familiarity and ease of use, the telephone may be appropriate to contact patients remotely. Individuals at risk of low eHealth literacy, such as older or less educated patients, may benefit from using a more traditional eHealth tool such as the telephone [53]. However, when comparing the effects on patient outcomes from studies using eHealth solutions other than the telephone as the delivery mode, telephone interventions do not stand out as more or less appropriate. We found that telephone interventions as a delivery mode effectively improved patients' self-care behaviors, but the effects on QoL and readmissions were less promising. This finding suggests that self-care follow-up is likely to be more important than the specific mode of follow-up.

Most studies in this review included features that required patients to monitor their vital signs and report health behaviors and symptoms. Giving patients with HF a more active role in their healing processes through posthospitalization eHealth interventions may promote their experience as true partners in shared decision-making, improve their well-being, and result in better adherence to treatment [54]. However, the value of eHealth interventions as part of health care for patients who are chronically ill may vary. Runz-Jørgensen et al [55] found that patients with multimorbidity and more significant illness and

treatment burden perceived eHealth interventions as more favorable than those with less complex disease and treatment. This result may be explained by considering the burdens of HF [55]. Patients with HF are vulnerable because they require regular and ongoing disease monitoring and management to reduce the risk of deterioration, and many fragile patients with HF have limited access to the health care system. The COVID-19 pandemic has forced health care systems to re-evaluate reimbursement for eHealth solutions to promote more widespread adoption of HF care [15,56,57].

We believe that for patients with HF to perceive the eHealth follow-up service as appropriate and be willing to use it, the timing of the introduction of the service is a crucial factor. In our review, patients were primarily enrolled in the eHealth interventions upon hospital discharge, ensuring patient support immediately after release. However, of the 18 included studies, 5 (28%) recruited patients during the first 4 weeks after discharge, demonstrating that eHealth interventions significantly increased QoL but had little impact on readmissions and self-care. Nevertheless, the findings of the effects from eHealth follow-up on patient outcomes suggest that patients with a severe heart condition benefit from prompt posthospitalization follow-up. It may be essential to provide patients with HF with self-care support at discharge to avoid 30-day readmission.

Remote monitoring as a feature of eHealth interventions may include parameters for detecting symptom and illness deterioration, successfully reducing readmissions among patients with HF [19-21]. However, in our study, the effects on readmissions from remote monitoring were inconclusive. For monitoring to be successful, aspects of measurement reliability and frequency, patient interface and adherence, and prompt interpretation by health professionals need to be considered [58]. Most of the included studies involving remote monitoring also provided contact with health care professionals, mainly nurses, who regularly stayed in touch with patients by either telephone calls or email. Koivunen and Soranto [59] identified communication and patient–nurse relationships as essential factors of telehealth in nursing practice. Patient–nurse interactions enable nurses to inquire about and assess patients' self-care needs and symptoms, express empathy, and increase patients' sense of security [58]. Another vital aspect of the patient–nurse interaction in the included eHealth interventions was whether the technology was acceptable to patients. Lack of required engagement among patients may be attributed to the nature of the technology [60], and patient adherence to the system is crucial for an intervention's success. Ding et al [61] found high adherence to the intervention component of weight monitoring (6 out of 7 days) in their recent telemonitoring RCT of patients with HF (published after our literature research). The intervention resulted in a significant improvement in self-management related to health maintenance, medication adherence, and diet [61]. We found that intervention adherence in most of the remote monitoring studies with patient–nurse interaction was 81%-99%. Comin-Colet et al [62] found that despite low expectations among patients before entering a telemedicine HF care intervention, adherence and satisfaction levels were high during the intervention, likely because of the HF care teams' proactive engagement with patients [62]. Clinicians who practice patient-centered communication adopting the patient's perspective may contribute to increased adherence levels in patients with HF, particularly during care transitions such as discharge from hospital to home, which to many patients can be confusing and demanding related to follow-up on treatment regimens [63]. The World Health Organization states that the quality of the treatment relationship is an essential determinant of adherence [26].

eHealth interventions have excellent potential to reinforce patient education on self-care [53]. Most of the reviewed studies in this review provided patients with education or counseling delivered by nurse specialists before the trial or during the trial, covering disease- and treatment-specific topics, psychosocial issues, and health behavior change. These studies seem to support improved self-care from eHealth interventions that include an educational aspect. Although the educational focus of many eHealth and mHealth interventions is illness management [64], a more holistic approach to self-care education not limited to only disease management is suggested. According to Lewis et al [65], addressing the holistic needs of patients with comorbidities using eHealth technology supports more patient-centered health care. Interestingly, of the 18 included studies, only 1 (6%) assessed changes in patients' knowledge at the completion of the intervention period. This study found that an HF education program involving iterative teaching tools expanded patients' HF knowledge [35]. This finding is in line with a review by Bashi et al [64], in which only 2 of the 15 mHealth interventions included an evaluation of patient knowledge as a study outcome. On the contrary, a recent Cochrane review of mHealth-delivered educational interventions for patients with HF found no evidence of a difference in HF knowledge or other patient-reported outcomes [66]. However, validated tools of patient knowledge can be an efficient measure of intervention success, and an assessment of patient knowledge as part of eHealth protocols is recommended [66].

Limitations

Several limitations should be mentioned. First, heterogeneity in the included studies made meta-analysis impossible, and a qualitative thematic analysis was applied. Such an analysis is prone to interpretation bias [33]. Second, the included eHealth interventions pertain to the transition phase between hospital and home, thus limiting generalizability to all stages of follow-up of patients with HF. Third, although most of the included studies indicated good methodological quality, most of them did not apply a blinded randomization process, and 50% (9/18) of the studies did not report adherence to the intervention.

Conclusions

This review identified 18 studies of posthospitalization follow-up interventions in patients with HF. Most of the included studies enrolled patients in eHealth interventions upon hospital discharge to ensure support in the critical post–hospital discharge period. The most common mode for posthospitalization follow-up was telemonitoring with telephone support. Patients received education or counseling about their disease, psychosocial issues, and health behavior changes. Most

studies also required the patients to monitor vital signs and report their health behaviors and symptoms.

The findings of the effects of interventions on patient outcomes such as QoL, self-care, and readmissions propose that patients with HF should receive prompt follow-up after hospital discharge. eHealth interventions, including patient education, support, and self-monitoring, have the potential to improve QoL, but it is less clear how eHealth interventions affect self-care behavior and readmissions in populations of patients with HF.

Aspects of measurement reliability and frequency, user interface and adherence, and prompt interpretation by health professionals need to be considered to ensure successful monitoring in eHealth interventions. These findings are important to inform future intervention studies to support patients with HF after discharge from the hospital. eHealth interventions have the potential to improve well-being, adherence to treatment, and patients' experiences of being engaged partners in shared decision-making.

Systematic reviews of the literature are recommended during the planning and development of complex interventions [29]. The findings from this review will be used to inform the development of a post–hospital discharge follow-up service addressing the burden of treatment and self-management among patients with HF.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Documentation of the literature search.
[PDF File (Adobe PDF File), 103 KB - jmir_v24i2e32946_app1.pdf]

References

1. GBD 2017 DiseaseInjury IncidencePrevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1789-1858 [FREE Full text] [doi: 10.1016/S0140-6736(18)32279-7] [Medline: 30496104]
2. Groenewegen A, Rutten FH, Mosterd A, Hoes AW. Epidemiology of heart failure. Eur J Heart Fail 2020 Aug;22(8):1342-1356 [FREE Full text] [doi: 10.1002/ejhf.1858] [Medline: 32483830]
3. Koudstaal S, Pujades-Rodriguez M, Denaxas S, Gho JM, Shah AD, Yu N, et al. Prognostic burden of heart failure recorded in primary care, acute hospital admissions, or both: a population-based linked electronic health record cohort study in 2.1 million people. Eur J Heart Fail 2017 Sep;19(9):1119-1127 [FREE Full text] [doi: 10.1002/ejhf.709] [Medline: 28008698]
4. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, ESC Scientific Document Group. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. Eur Heart J 2021 Sep 21;42(36):3599-3726. [doi: 10.1093/eurheartj/ehab368] [Medline: 34447992]
5. Nordfonn OK, Morken IM, Bru LE, Larsen AI, Husebø AM. Burden of treatment in patients with chronic heart failure - A cross-sectional study. Heart Lung 2021;50(3):369-374 [FREE Full text] [doi: 10.1016/j.hrtlng.2021.02.003] [Medline: 33618147]
6. Khan MS, Sreenivasan J, Lateef N, Abougergi MS, Greene SJ, Ahmad T, et al. Trends in 30- and 90-day readmission rates for heart failure. Circ Heart Fail 2021 Apr;14(4):e008335. [doi: 10.1161/CIRCHEARTFAILURE.121.008335] [Medline: 33866827]
7. Albert NM, Barnason S, Deswal A, Hernandez A, Kociol R, Lee E, American Heart Association Complex Cardiovascular PatientFamily Care Committee of the Council on CardiovascularStroke Nursing, Council on Clinical Cardiology,Council on Quality of CareOutcomes Research. Transitions of care in heart failure: a scientific statement from the American Heart Association. Circ Heart Fail 2015 Mar;8(2):384-409. [doi: 10.1161/HHF.0000000000000006] [Medline: 25604605]
8. Cella DF. Quality of life: concepts and definition. J Pain Symptom Manage 1994 Apr;9(3):186-192 [FREE Full text] [doi: 10.1016/0885-3924(94)90129-5] [Medline: 8014530]
9. Lee CS, Bidwell JT, Paturzo M, Alvaro R, Cocchieri A, Jaarsma T, et al. Patterns of self-care and clinical events in a cohort of adults with heart failure: 1 year follow-up. Heart Lung 2018;47(1):40-46 [FREE Full text] [doi: 10.1016/j.hrtlng.2017.09.004] [Medline: 29054487]
10. Jaarsma T, Hill L, Bayes-Genis A, La Rocca HB, Castiello T, Čelutkienė J, et al. Self-care of heart failure patients: practical management recommendations from the Heart Failure Association of the European Society of Cardiology. Eur J Heart Fail 2021 Jan;23(1):157-174 [FREE Full text] [doi: 10.1002/ejhf.2008] [Medline: 32945600]
11. Pollack AH, Backonja U, Miller AD, Mishra SR, Khelifi M, Kendall L, et al. Closing the gap: supporting patients' transition to self-management after hospitalization. Proc SIGCHI Conf Hum Factor Comput Syst 2016 May;2016:5324-5336 [FREE Full text] [doi: 10.1145/2858036.2858240] [Medline: 27500285]
12. Gracia E, Singh P, Collins S, Chioncel O, Pang P, Butler J. The vulnerable phase of heart failure. Am J Ther 2018;25(4):e456-e464 [FREE Full text] [doi: 10.1097/MJT.0000000000000794] [Medline: 29985824]

13. Takeda A, Taylor SJ, Taylor RS, Khan F, Krum H, Underwood M. Clinical service organisation for heart failure. Cochrane Database Syst Rev 2012 Sep 12(9):CD002752. [doi: 10.1002/14651858.CD002752.pub3] [Medline: 22972058]
14. Lee KK, Yang J, Hernandez AF, Steimle AE, Go AS. Post-discharge follow-up characteristics associated with 30-day readmission after heart failure hospitalization. Med Care 2016 Apr;54(4):365-372 [FREE Full text] [doi: 10.1097/MLR.0000000000000492] [Medline: 26978568]
15. Ski CF, Zippel-Schultz B, De Maesschalck L, Hoedemakers T, Schütt K, Thompson DR, et al. COVID-19 shapes the future for management of patients with chronic cardiac conditions. Digit Health 2021;7:2055207621991711 [FREE Full text] [doi: 10.1177/2055207621991711] [Medline: 33623705]
16. Storm M, Siemsen IM, Laugaland K, Dyrstad DN, Aase K. Quality in transitional care of the elderly: key challenges and relevant improvement measures. Int J Integr Care 2014 Apr;14:e013 [FREE Full text] [doi: 10.5334/ijic.1194] [Medline: 24868196]
17. Eysenbach G. What is e-health? J Med Internet Res 2001;3(2):E20 [FREE Full text] [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
18. Usak M, Kubiatko M, Shabbir M, Viktorovna Dudnik O, Jermsittiparsert K, Rajabion L. Health care service delivery based on the internet of things: a systematic and comprehensive study. Int J Commun Syst 2019 Sep 13;33(2):e4179 [FREE Full text] [doi: 10.1002/dac.4179]
19. Auener SL, Remers TE, van Dulmen SA, Westert GP, Kool RB, Jeurissen PP. The effect of noninvasive telemonitoring for chronic heart failure on health care utilization: systematic review. J Med Internet Res 2021 Sep 29;23(9):e26744 [FREE Full text] [doi: 10.2196/26744] [Medline: 34586072]
20. Ding H, Chen SH, Edwards I, Jayasena R, Doecke J, Layland J, et al. Effects of different telemonitoring strategies on chronic heart failure care: systematic review and subgroup meta-analysis. J Med Internet Res 2020 Nov 13;22(11):e20032 [FREE Full text] [doi: 10.2196/20032] [Medline: 33185554]
21. Lin M, Yuan W, Huang T, Zhang H, Mai J, Wang J. Clinical effectiveness of telemedicine for chronic heart failure: a systematic review and meta-analysis. J Investig Med 2017 Jun;65(5):899-911. [doi: 10.1136/jim-2016-000199] [Medline: 28330835]
22. Rosen O, Fridman R, Rosen B, Shane R, Pevnick J. Medication adherence as a predictor of 30-day hospital readmissions. Patient Prefer Adherence 2017 Apr 20;11:801-810 [FREE Full text] [doi: 10.2147/PPA.S125672] [Medline: 28461742]
23. Ma C. Rehospitalisation rates and associated factors within 6 months after hospital discharge for patients with chronic heart failure: a longitudinal observational study. J Clin Nurs 2019 Jul;28(13-14):2526-2536. [doi: 10.1111/jocn.14830] [Medline: 30793392]
24. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015 Feb 24;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
25. Fernandez-Lazaro CI, García-González JM, Adams DP, Fernandez-Lazaro D, Mielgo-Ayuso J, Caballero-Garcia A, et al. Adherence to treatment and related factors among patients with chronic conditions in primary care: a cross-sectional study. BMC Fam Pract 2019 Sep 14;20(1):132 [FREE Full text] [doi: 10.1186/s12875-019-1019-3] [Medline: 31521114]
26. De Geest S, Sabaté E. Adherence to long-term therapies: evidence for action. Eur J Cardiovasc Nurs 2003 Dec;2(4):323. [Medline: 14667488]
27. Hulley SB, Newman TB, Cummings SR. Getting started: the anatomy and physiology of clinical research. In: Hulley SB, Cummings SR, Browner WS, Grady DE, Newman TB, editors. Designing Clinical Research. Philadelphia: Lippincott Williams & Wilkins; 2007.
28. Plüddemann A, Aronson JK, Onakpoya I, Heneghan C, Mahtani KR. Redefining rapid reviews: a flexible framework for restricted systematic reviews. BMJ Evid Based Med 2018 Dec;23(6):201-203. [doi: 10.1136/bmjebm-2018-110990] [Medline: 29950313]
29. Skivington K, Matthews L, Craig P, Simpson S, Moore L. Developing and evaluating complex interventions: updating Medical Research Council guidance to take account of new methodological and theoretical approaches. The Lancet 2018 Nov;392:S2. [doi: 10.1016/s0140-6736(18)32865-4]
30. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. Cochrane Handbook for Systematic Reviews of Interventions, 2nd edition. Chichester, UK: John Wiley & Sons; 2019. [doi: 10.1002/9781119536604]
31. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
32. Hong QN, Gonzalez-Reyes A, Pluye P. Improving the usefulness of a tool for appraising the quality of qualitative, quantitative and mixed methods studies, the Mixed Methods Appraisal Tool (MMAT). J Eval Clin Pract 2018 Jun 21;24(3):459-467. [doi: 10.1111/jep.12884] [Medline: 29464873]
33. Whittemore R, Knafl K. The integrative review: updated methodology. J Adv Nurs 2005 Dec;52(5):546-553. [doi: 10.1111/j.1365-2648.2005.03621.x] [Medline: 16268861]
34. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 05;5(1):210 [FREE Full text] [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]

35. Athilingam P, Jenkins B, Johansson M, Labrador M. A mobile health intervention to improve self-care in patients with heart failure: pilot randomized control trial. JMIR Cardio 2017 Aug 11;1(2):e3 [FREE Full text] [doi: 10.2196/cardio.7848] [Medline: 31758759]
36. Dunbar SB, Reilly CM, Gary R, Higgins MK, Culler S, Butts B, et al. Randomized clinical trial of an integrated self-care intervention for persons with heart failure and diabetes: quality of life and physical functioning outcomes. J Card Fail 2015 Sep;21(9):719-729 [FREE Full text] [doi: 10.1016/j.cardfail.2015.05.012] [Medline: 26028261]
37. Evangelista L, Lee J, Moore A, Motie M, Ghasemzadeh H, Sarrafzadeh M, et al. Examining the effects of remote monitoring systems on activation, self-care, and quality of life in older patients with chronic heart failure. J Cardiovasc Nurs 2015;30(1):51-57 [FREE Full text] [doi: 10.1097/JCN.0000000000000110] [Medline: 24365871]
38. Gallagher BD, Moise N, Haerizadeh M, Ye S, Medina V, Kronish IM. Telemonitoring Adherence to Medications in Heart Failure Patients (TEAM-HF): a pilot randomized clinical trial. J Card Fail 2017 Apr;23(4):345-349 [FREE Full text] [doi: 10.1016/j.cardfail.2016.11.001] [Medline: 27818309]
39. Jayaram NM, Khariton Y, Krumholz HM, Chaudhry SI, Mattera J, Tang F, et al. Impact of telemonitoring on health status. Circ Cardiovasc Qual Outcomes 2017 Dec;10(12)::e004148 [FREE Full text] [doi: 10.1161/CIRCOUTCOMES.117.004148] [Medline: 29237746]
40. Masterson Creber R, Patey M, Lee CS, Kuan A, Jurgens C, Riegel B. Motivational interviewing to improve self-care for patients with chronic heart failure: MITI-HF randomized controlled trial. Patient Educ Couns 2016 Feb;99(2):256-264 [FREE Full text] [doi: 10.1016/j.pec.2015.08.031] [Medline: 26358533]
41. Ong MK, Romano PS, Edgington S, Aronow HU, Auerbach AD, Black JT, Better Effectiveness After Transition–Heart Failure (BEAT-HF) Research Group. Effectiveness of remote patient monitoring after discharge of hospitalized patients with heart failure: the Better Effectiveness After Transition -- Heart Failure (BEAT-HF) randomized clinical trial. JAMA Intern Med 2016 Mar;176(3):310-318 [FREE Full text] [doi: 10.1001/jamainternmed.2015.7712] [Medline: 26857383]
42. Ritchie CS, Houston TK, Richman JS, Sobko HJ, Berner ES, Taylor BB, et al. The E-Coach technology-assisted care transition system: a pragmatic randomized trial. Transl Behav Med 2016 Sep;6(3):428-437 [FREE Full text] [doi: 10.1007/s13142-016-0422-8] [Medline: 27339715]
43. Srivastava A, Do J, Sales VL, Ly S, Joseph J. Impact of patient-centred home telehealth programme on outcomes in heart failure. J Telemed Telecare 2019 Aug;25(7):425-430. [doi: 10.1177/1357633X18775852] [Medline: 29793388]
44. Young L, Hertzog M, Barnason S. Effects of a home-based activation intervention on self-management adherence and readmission in rural heart failure patients: the PATCH randomized controlled trial. BMC Cardiovasc Disord 2016 Sep 08;16(1):176 [FREE Full text] [doi: 10.1186/s12872-016-0339-7] [Medline: 27608624]
45. Comín-Colet J, Enjuanes C, Verdú-Rotellar JM, Linas A, Ruiz-Rodriguez P, González-Robledo G, et al. Impact on clinical events and healthcare costs of adding telemedicine to multidisciplinary disease management programmes for heart failure: results of a randomized controlled trial. J Telemed Telecare 2016 Jul;22(5):282-295. [doi: 10.1177/1357633X15600583] [Medline: 26350543]
46. Frederix I, Vanderlinden L, Verboven A, Welten M, Wouters D, De Keulenaer G, et al. Long-term impact of a six-month telemedical care programme on mortality, heart failure readmissions and healthcare costs in patients with chronic heart failure. J Telemed Telecare 2019 Jun;25(5):286-293. [doi: 10.1177/1357633X18774632] [Medline: 29742959]
47. Kotooka N, Kitakaze M, Nagashima K, Asaka M, Kinugasa Y, Nochioka K, HOMES-HF study investigators. The first multicenter, randomized, controlled trial of home telemonitoring for Japanese patients with heart failure: home telemonitoring study for patients with heart failure (HOMES-HF). Heart Vessels 2018 Aug;33(8):866-876. [doi: 10.1007/s00380-018-1133-5] [Medline: 29450689]
48. Hwang R, Bruning J, Morris NR, Mandrusiak A, Russell T. Home-based telerehabilitation is not inferior to a centre-based program in patients with chronic heart failure: a randomised trial. J Physiother 2017 Apr;63(2):101-107 [FREE Full text] [doi: 10.1016/j.jphys.2017.02.017] [Medline: 28336297]
49. Lycholip E, Thon Aamodt I, Lie I, Šimbelytė T, Puronaitė R, Hillege H, et al. The dynamics of self-care in the course of heart failure management: data from the IN TOUCH study. Patient Prefer Adherence 2018;12:1113-1122 [FREE Full text] [doi: 10.2147/PPA.S162219] [Medline: 29983549]
50. Kraai I, de Vries A, Vermeulen K, van Deursen V, van der Wal M, de Jong R, et al. The value of telemonitoring and ICT-guided disease management in heart failure: results from the IN TOUCH study. Int J Med Inform 2016 Jan;85(1):53-60 [FREE Full text] [doi: 10.1016/j.ijmedinf.2015.10.001] [Medline: 26514079]
51. Köberich S, Lohrmann C, Mittag O, Dassen T. Effects of a hospital-based education programme on self-care behaviour, care dependency and quality of life in patients with heart failure--a randomised controlled trial. J Clin Nurs 2015 Jun;24(11-12):1643-1655. [doi: 10.1111/jocn.12766] [Medline: 25661470]
52. Pedone C, Rossi FF, Cecere A, Costanzo L, Antonelli Incalzi R. Efficacy of a physician-led multiparametric telemonitoring system in very old adults with heart failure. J Am Geriatr Soc 2015 Jun;63(6):1175-1180. [doi: 10.1111/jgs.13432] [Medline: 26031737]
53. Levin-Zamir D, Bertschi I. Media health literacy, eHealth literacy, and the role of the social environment in context. Int J Environ Res Public Health 2018 Aug 03;15(8):1643 [FREE Full text] [doi: 10.3390/ijerph15081643] [Medline: 30081465]

54. Epstein RM, Fiscella K, Lesser CS, Stange KC. Why the nation needs a policy push on patient-centered health care. Health Aff (Millwood) 2010 Aug;29(8):1489-1495. [doi: 10.1377/hlthaff.2009.0888] [Medline: 20679652]
55. Runz-Jørgensen SM, Schiøtz ML, Christensen U. Perceived value of eHealth among people living with multimorbidity: a qualitative study. J Comorb 2017;7(1):96-111 [FREE Full text] [doi: 10.15256/joc.2017.7.98] [Medline: 29359124]
56. Singhal A, Cowie MR. Digital health: implications for heart failure management. Card Fail Rev 2021 Mar;7:e08 [FREE Full text] [doi: 10.15420/cfr.2020.28] [Medline: 34035953]
57. Ansari Ramandi MM, Yarmohammari H, Naderi N. A call for telehealth application for heart failure patients during the Coronavirus 2019 pandemic. ESC Heart Fail 2021 Aug;8(4):3431-3432 [FREE Full text] [doi: 10.1002/ehf2.13387] [Medline: 33934547]
58. Mataxen PA, Webb LD. Telehealth nursing: more than just a phone call. Nursing 2019 Apr;49(4):11-13. [doi: 10.1097/01.NURSE.0000553272.16933.4b] [Medline: 30893196]
59. Koivunen M, Saranto K. Nursing professionals' experiences of the facilitators and barriers to the use of telehealth applications: a systematic review of qualitative studies. Scand J Caring Sci 2018 Mar;32(1):24-44. [doi: 10.1111/scs.12445] [Medline: 28771752]
60. Inglis SC, Clark RA, McAlister FA, Stewart S, Cleland JG. Which components of heart failure programmes are effective? A systematic review and meta-analysis of the outcomes of structured telephone support or telemonitoring as the primary component of chronic heart failure management in 8323 patients: abridged Cochrane review. Eur J Heart Fail 2011 Sep;13(9):1028-1040 [FREE Full text] [doi: 10.1093/eurjhf/hfr039] [Medline: 21733889]
61. Ding H, Jayasena R, Chen SH, Maiorana A, Dowling A, Layland J, et al. The effects of telemonitoring on patient compliance with self-management recommendations and outcomes of the innovative telemonitoring enhanced care program for chronic heart failure: randomized controlled trial. J Med Internet Res 2020 Jul 08;22(7):e17559 [FREE Full text] [doi: 10.2196/17559] [Medline: 32673222]
62. Comín-Colet J, Enjuanes C, Lupón J, Cainzos-Achirica M, Badosa N, Verdú JM. Transitions of care between acute and chronic heart failure: critical steps in the design of a multidisciplinary care model for the prevention of rehospitalization. Rev Esp Cardiol (Engl Ed) 2016 Oct;69(10):951-961. [doi: 10.1016/j.rec.2016.05.001] [Medline: 27282437]
63. Piña IL, Di Palo KE, Brown MT, Choudhry NK, Cvengros J, Whalen D, et al. Medication adherence: importance, issues and policy: a policy statement from the American Heart Association. Prog Cardiovasc Dis 2021;64:111-120. [doi: 10.1016/j.pcad.2020.08.003] [Medline: 32800791]
64. Bashi N, Fatehi F, Fallah M, Walters D, Karunanithi M. Self-management education through mHealth: review of strategies and structures. JMIR Mhealth Uhealth 2018 Oct 19;6(10):e10771 [FREE Full text] [doi: 10.2196/10771] [Medline: 30341042]
65. Lewis J, Ray P, Liaw S. Recent worldwide developments in eHealth and mHealth to more effectively manage cancer and other chronic diseases - a systematic review. Yearb Med Inform 2016 Nov 10(1):93-108 [FREE Full text] [doi: 10.15265/IY-2016-020] [Medline: 27830236]
66. Allida S, Du H, Xu X, Prichard R, Chang S, Hickman LD, et al. mHealth education interventions in heart failure. Cochrane Database Syst Rev 2020 Jul 02;7:CD011845 [FREE Full text] [doi: 10.1002/14651858.CD011845.pub2] [Medline: 32613635]

Abbreviations

HF: heart failure
mHealth: mobile health
MMAT: Mixed Methods Appraisal Tool
QoL: quality of life
RCT: randomized controlled trial

Edited by R Kukafka; submitted 16.08.21; peer-reviewed by S Ye, M Johansson; comments to author 15.09.21; revised version received 09.11.21; accepted 03.12.21; published 15.02.22.

Please cite as:
Morken IM, Storm M, Søreide JA, Urstad KH, Karlsen B, Husebø AML
Posthospitalization Follow-Up of Patients With Heart Failure Using eHealth Solutions: Restricted Systematic Review
J Med Internet Res 2022;24(2):e32946
URL: https://www.jmir.org/2022/2/e32946
doi:10.2196/32946
PMID:35166680

Review

Remotely Delivered Interventions to Support Women With Symptoms of Anxiety in Pregnancy: Mixed Methods Systematic Review and Meta-analysis

Kerry Evans[1*], BSc, MA, PhD; Stefan Rennick-Egglestone[2*], BA, PhD; Serena Cox[1*], BSc, MSc; Yvonne Kuipers[3*], PhD; Helen Spiby[1*], MPhil

[1]School of Health Sciences, University of Nottingham, Nottingham, United Kingdom

[2]Institute of Mental Health, School of Health Sciences, University of Nottingham, Nottingham, United Kingdom

[3]School of Health and Social Care, Edinburgh Napier University, Edinburgh, United Kingdom

[*]all authors contributed equally

Corresponding Author:
Kerry Evans, BSc, MA, PhD
School of Health Sciences
University of Nottingham
Queen's Medical Centre
Derby Rd, Lenton
Nottingham, NG7 2HA
United Kingdom
Phone: 44 7596783920
Email: kerry.evans1@nottingham.ac.uk

Abstract

Background: Symptoms of anxiety are common in pregnancy, with severe symptoms associated with negative outcomes for women and babies. Low-level psychological therapy is recommended for women with mild to moderate anxiety, with the aim of preventing an escalation of symptoms and providing coping strategies. Remotely delivered interventions have been suggested to improve access to treatment and support and provide a cost-effective, flexible, and timely solution.

Objective: This study identifies and evaluates remotely delivered, digital, or web-based interventions to support women with symptoms of anxiety during pregnancy.

Methods: This mixed methods systematic review followed a convergent segregated approach to synthesize qualitative and quantitative data. The ACM Digital Library, Allied and Complementary Medicine Database, Applied Social Sciences Index and Abstracts, Centre for Reviews and Dissemination database, the Cochrane Central Register of Controlled Trials, the Cochrane Library, CINAHL, Embase, Health Technology Assessment Library, IEEE Xplore, Joanna Briggs Institute, Maternity and Infant Care, MEDLINE, PsycINFO, and the Social Science Citation Index were searched in October 2020. Quantitative or qualitative primary research that included pregnant women and evaluated remotely delivered interventions reporting measures of anxiety, fear, stress, distress, women's views, and opinions were included.

Results: Overall, 3 qualitative studies and 14 quantitative studies were included. Populations included a general antenatal population and pregnant women having anxiety and depression, fear of childbirth, insomnia, and preterm labor. Interventions included cognitive behavioral therapy, problem solving, mindfulness, and educational designs. Most interventions were delivered via web-based platforms, and 62% (8/13) included direct contact from trained therapists or coaches. A meta-analysis of the quantitative data found internet-based cognitive behavioral therapy and facilitated interventions showed a beneficial effect in relation to the reduction of anxiety scores (standardized mean difference −0.49, 95% CI −0.75 to −0.22; standardized mean difference −0.48, 95% CI −0.75 to −0.22). Due to limitations in the amount of available data and study quality, the findings should be interpreted with caution. Synthesized findings found some evidence to suggest that interventions are more effective when women maintain regular participation which may be enhanced by providing regular contact with therapists or peer support, appropriate targeting of interventions involving components of relaxation and cognitive-based skills, and providing sufficient sessions to develop new skills without being too time consuming.

Conclusions: There is limited evidence to suggest that women who are pregnant may benefit from remotely delivered interventions. Components of interventions that may improve the effectiveness and acceptability of remotely delivered interventions included

providing web-based contact with a therapist, health care professional, or peer community. Women may be more motivated to complete interventions that are perceived as relevant or tailored to their needs. Remote interventions may also provide women with greater anonymity to help them feel more confident in disclosing their symptoms.

(J Med Internet Res 2022;24(2):e28093) doi:10.2196/28093

KEYWORDS

anxiety; pregnancy; antenatal; systematic review; digital interventions; eHealth; remote interventions; electronic health; parenting; remote delivery; therapy; CBT; fear; distress; mobile phone

Introduction

Symptoms of anxiety are common in pregnancy and have been reported across all 3 trimesters [1,2]. Recent reviews have reported worldwide prevalence rates of 15% to 20% in the antenatal period, with higher rates reported in low- to middle-income countries, particularly among poorer women with gender-based or psychiatric risks [3-5]. Risk factors for developing anxiety during pregnancy include demographic and socioeconomic factors (lack of partner support, young age, lower education, smoking, and being overweight), psychological factors (past mental health concerns, reduced emotional well-being, stressful events, low self-esteem, and negative coping styles), and obstetric factors (previous pregnancy loss and complications in pregnancy) [6,7]. Antenatal anxiety is reported to be associated with postpartum depression, reduced rates of breastfeeding, prematurity, and preterm birth [8,9].

Facilitated self-help and low-intensity psychological interventions are recommended as the first-line treatment option in pregnancy within a stepped care approach [10]; however, there are limited data on the effectiveness of interventions for the antenatal period [11]. Vigod and Dennis [12] discussed that interventions delivered via the internet might present a solution for overcoming the barriers of access to treatment for perinatal mental health disorders, such as a lack of specialist psychological and psychiatric support, the cost of services and transportation, and childcare requirements. Interventions can be delivered as unguided resources to support or replace patient–provider interactions or as guided interventions that may include live interactions over telephone or video or contact with therapists using digital messaging. Web-based interventions and peer discussion groups may benefit women by reducing the stigma of mental illness in pregnancy, addressing treatment barriers, and strengthening social support mechanisms. Web-based therapist-assisted interventions may offer women flexibility and convenience and be more efficient than the delivery of one-on-one therapy [12]. In particular, the use of mobile technology is thought to offer a low-cost solution for improving treatment accessibility and sustainability. In high-income countries, between 84% and 99% of people aged <35 years own a smartphone, which would include women in their prime reproductive years [13,14].

A recent review of internet-delivered psychological interventions for perinatal anxiety and depression [15] included 2 cognitive behavioral therapy (CBT) interventions for women with fear of childbirth [16] and symptoms of anxiety and depression [17]. Improvement in symptoms was reported, and participant satisfaction was positive in both studies. A scoping review of mobile health apps and text-based interventions for perinatal anxiety and depression (n=26 publications from 22 studies) [14] identified 1 intervention focused solely on anxiety and 9 interventions for both depression and anxiety. Intervention strategies included peer support, psychoeducation, and active therapy.

Anxiety symptom screening for maternity care, which, in the United Kingdom, uses a simple 2-item measure (Generalized Anxiety Disorder [GAD] 2 items) [18], is recommended to identify where women may need further support or referral for specific diagnosis by specialist mental health professionals. Locating interventions that can be integrated within current maternity care structures was a key motivation in the design of this review. Maternity care providers require guidance on the aspects of care that can be safely and effectively delivered via remote interventions and consider the implications in terms of acceptability, fidelity, and equitability [19-21]. The COVID-19 pandemic presented a further motivation for this review, considering the impact of the pandemic on women's mental health and identifying innovative ways of delivering safe, effective, and equitable care [22,23]. In this review, we seek to identify the types of remotely delivered interventions that are available and have been evaluated to improve the symptoms of anxiety in women who are pregnant. We include a broad concept of common anxiety disorders in pregnancy, including pregnancy-specific anxiety and fear of birth, and a range of intervention strategies, including psychological, mind–body, education, and social support. The review aims to answer the following questions:

- What remotely delivered, digital, or web-based interventions have been evaluated and reported in the research literature to improve the symptoms of anxiety in women who are pregnant?
- How effective are remotely delivered, digital, or web-based interventions in reducing the symptoms of anxiety in pregnancy?
- What are women's views, attitudes, and experiences of accessing and participating in remotely delivered anxiety interventions during pregnancy?

Methods

Overview

A mixed methods systematic review was conducted following the established Joanna Briggs Institute (JBI) methodological guidance [24]. This approach supported the broad research aim and enabled evidence from diverse study designs to provide a comprehensive and detailed understanding of the current

evidence base [25,26]. The review adopted a convergent segregated mixed methods approach [27] in which the search and synthesis process for the different study designs included in the review were conducted concurrently. Qualitative and quantitative evidence were initially analyzed separately, followed by comparison and integration of qualitative and quantitative data to form an overarching synthesis. Before commencement, the review protocol was registered on the PROSPERO (The International Prospective Register of Systematic Reviews) database (CRD42020195887).

Eligibility Criteria

Papers were included if the following criteria were met:

1. Participants included women who are pregnant of all parities across the 3 trimesters of pregnancy; women who are pregnant and under the care of specialist mental health services for severe and enduring mental health conditions were excluded
2. Evaluated the following types of interventions: mind–body interventions (relaxation, yoga, meditation, mindfulness, hypnotherapy, and imagery), social support interventions (supportive interactions, group discussions, peer support, telephone support, and exercise), educational interventions (birth preparation, educational sessions, educational materials, psychoeducational interventions, and antenatal classes), psychological interventions (CBT, mindfulness CBT, group CBT, interpersonal psychological therapy, nondirective counseling, therapy, and problem-solving therapy), and interventions that have been remotely delivered, including web-based materials, via telephone, computer software, digital apps, and digital forums
3. Comparators or control groups (CGs) included usual antenatal care or other types of interventions
4. Interventions were considered appropriate and capable of being introduced into maternity care
5. Outcome evaluations included measures of anxiety, fear, stress, distress, and women's views, feedback, and opinions
6. Study methods included quantitative or qualitative primary research, including mixed methods studies involving any number of participants

Information Sources and Search Strategy

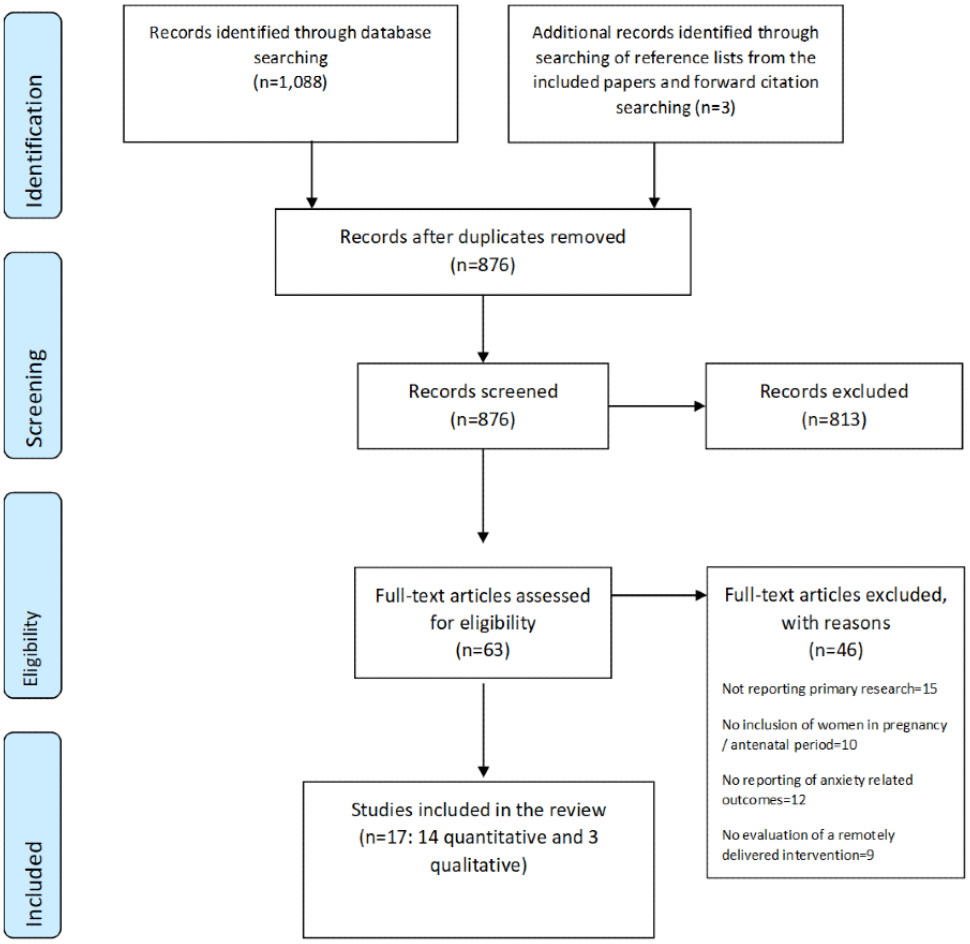
A search of the following electronic bibliographic databases was undertaken (October to November 2020): the ACM Digital Library; Allied and Complementary Medicine Database; Applied Social Sciences Index and Abstracts; Centre for Reviews and Dissemination database; the Cochrane Central Register of Controlled Trials; the Cochrane Library; the Cochrane Depression, Anxiety and Neurosis Group's Trials Register; CINAHL; Embase; Health Technology Assessment Library; IEEE Xplore; JBI; Maternity and Infant Care; MEDLINE; PsycINFO; and the Social Science Citation Index. References cited in existing systematic reviews and meta-analyses and reference lists of identified studies were searched to identify additional potentially relevant studies.

The search was limited to studies published since 2000 focusing on women who are pregnant and written in English. This reflects the period in which digitally delivered interventions were available in maternity care.

Titles and abstracts of papers were screened independently by 2 reviewers (KE and SC) against the inclusion and exclusion criteria to identify potentially relevant papers. Potentially relevant papers were retrieved and read in full to identify papers for inclusion in the systematic review. The papers identified for inclusion were agreed upon following discussions with the entire review team. A PRISMA (Preferred Reporting Item for Systematic Reviews and Meta-Analyses) flowchart (Figure 1) [28] was completed to summarize the study selection process. Reference management (RevMan, Version 5.4; The Cochrane Collaboration, 2020) software was used to organize and catalog the references.

Search terms used in the review are summarized in Textbox 1.

Figure 1. PRISMA (Preferred Reporting Item for Systematic Reviews and Meta-Analyses) diagram, remotely delivered interventions for anxiety in pregnancy.



Textbox 1. Summary of search terms.

Search terms

- "exp Anxiety/"
- "Anxiety Disorders/"
- "Fear/"
- "worries.mp"
- "worry.mp"
- "anxiety.mp"
- "anxious.mp"
- "fear.mp"

AND

- "Computer assisted instruction/"
- "Software/"
- "Social Media/"
- "Computer Communication Networks/"
- "Internet/"
- "Telemedicine/Telehealth/"
- "Internet-based intervention/"
- "Online intervention/"
- "Mobile applications/"
- "Information services/"
- "Video games/"
- "Monitoring physiologic/"
- "digital.mp"
- "electronic.mp"
- "e-learning.mp"
- "mobile applications.mp"
- "web.mp"
- "web-based.mp"
- "virtual.mp"
- "telephone.mp"
- "skype.mp"
- "face-time.mp"
- "video.mp"
- "video conferenceing.mp"

AND

- "Pregnancy/"
- "Peripartum period/"
- "Perinatal Care/"
- "childbearning.mp"
- "ante*natal.mp"
- "ante*partum.mp"
- "pregnancy.mp"

- "peripartum.mp"
- "perinatal.mp"

AND

- "Clinical Trial/"
- "Randomized Controlled Trial/"
- "systematic review"/
- "Meta-analysis/"
- "Cohort studies/"
- "Non-randomised/ Prospective/"
- "Case-Control studies/"
- "Qualitative research/"
- "Research Design/"
- "cohort.mp"
- "rct.mp"
- "randomised controlled trial.mp"
- "Quasi-experimental.mp"
- "Intervention.mp"
- "Case-control.mp"
- "Focus groups/"
- "Interviews/"
- "User experience.mp"
- "Design research.mp"
- "Intervention.mp"

Risk of Bias and Quality Assessment

The quantitative experimental study designs were independently assessed by 2 reviewers using the Cochrane Collaboration Effective Practice and Organization of Care tool for assessing the risk of bias [29]. This assesses the risk of selection, performance, detection, attrition, and reporting biases. The RevMan (version 5) software package was used to assist the organization and presentation of the data. The Critical Appraisal Skills Program for qualitative and cohort studies, the Critical Appraisal Skills Program for qualitative and cohort studies [30], The National Heart, Lung, and Blood Institute Quality Assessment Tool for Before–After (Pre–Post) Studies with No Control Group [31], and the JBI Checklist for Quasi-Experimental Studies [32] were also used to assist the methodological assessment of studies for inclusion in the systematic review analysis. Although no studies were excluded on the basis of quality, the quality assessment was used to identify the strengths and limitations of the review [24].

Data Extraction

Data extraction forms were designed and piloted. Extraction was completed by 2 independent researchers (KE and SC). Data extraction tables comprising numerical and textual data were produced to present the study characteristics, results, and quality assessments.

Data Synthesis

Tables were produced to present the study characteristics, results of interest, risk of bias ratings for randomized controlled trials (RCTs), and quality ratings for the included papers. This is presented alongside a narrative description of the data. A meta-analysis of quantitative data from the included experimental studies was conducted [29]. Considerations for performing meta-analysis included assessing the risk of bias of the studies and performing a qualitative assessment of clinical homogeneity. To evaluate statistical heterogeneity, the Q statistic and I^2 index were used [33]. A random effects model was identified as the most appropriate method of analysis, and subgroup analysis was planned for subsets of intervention designs, participants, and methods of delivery. The standardized mean difference (SMD) was presented as the summary statistic for continuous data for anxiety and other outcomes, with 95% CIs and 2-sided *P* values calculated for each outcome where possible. A Grading of Recommendations Assessment, Development and Evaluation (GRADE) assessment [29,34] of the quantitative review findings was conducted, and a GRADE Confidence in Evidence from Reviews of Qualitative research (GRADE-CERQual) assessment of the qualitative studies [35] was planned, although there were insufficient data to perform a GRADE-CERQual assessment.

Following the convergent segregated approach [27], quantitative findings were translated into narrative statements that were then compared, contrasted, and juxtaposed against the qualitative findings and integrated to configure synthesized findings to address the research questions. The synthesized findings were then interpreted and configured into overarching themes to develop explanation, understanding, and coherence between the synthesized findings as a whole [26,36].

Results

Details of the included qualitative and quantitative studies, population, intervention components, and delivery are presented in Multimedia Appendix 1 (TIDieR [Template for Intervention Description and Replication] checklist) [37].

Quality Assessment

The risk of bias assessments for quantitative experimental studies are presented in Figures 2 and 3. Studies were assessed as having a high or moderate risk of bias against the risk of bias standards for studies with a separate CG [29]. A cohort feasibility study [16] was assessed as moderate for methodological quality, which was mainly attributed to a lack of reporting of the recruitment strategy and the small sample size. The 3 qualitative studies were assessed using a moderate to high methodological quality score. Moderate scores were allocated to studies assessed as having limited reporting of (1) researchers' reflexivity and (2) Sampling and context [38,39].

Figure 2. Risk of bias table with assessments of the experimental studies [16,17,38-41,43-47,49-52,62].

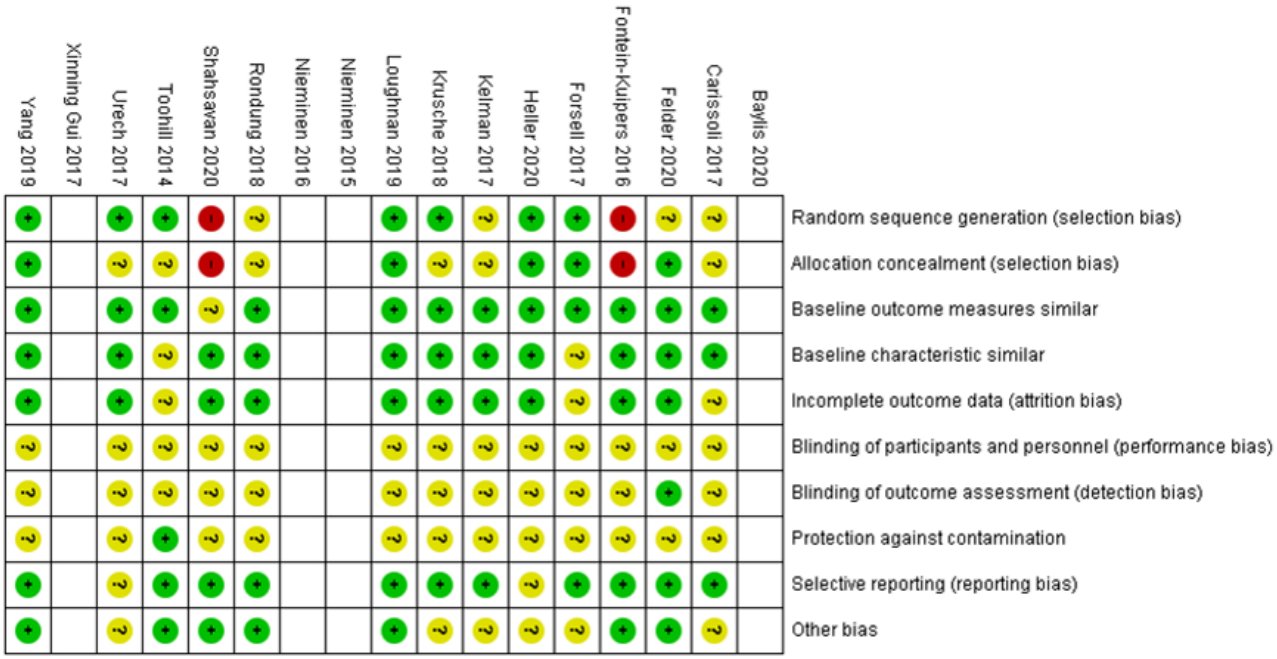
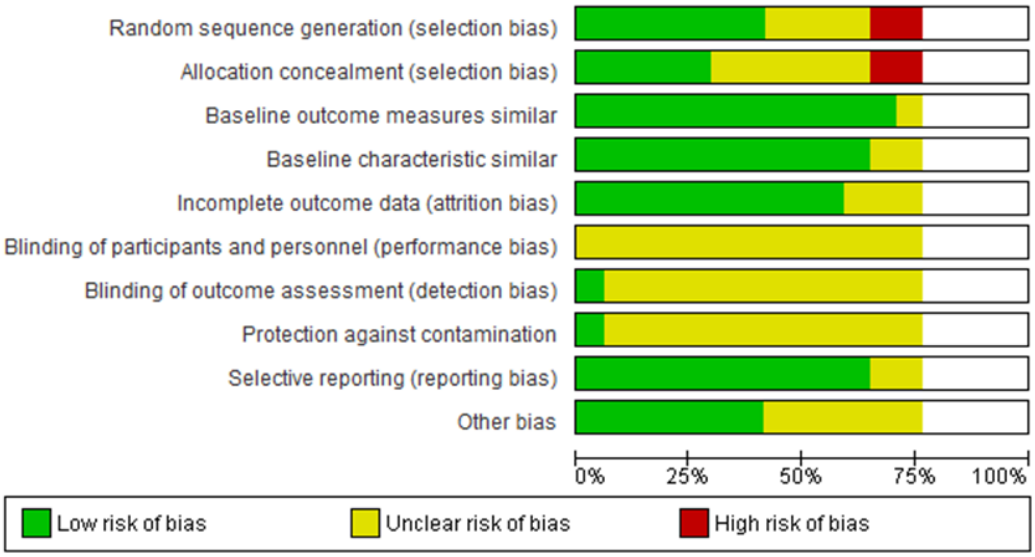


Figure 3. Risk of bias summary of the quantitative experimental selected studies.



Results From the Quantitative Studies

Overview

Of the 16 included studies, 14 (88%) were conducted in Australia, China, Iran, Italy, the Netherlands, Sweden, Switzerland, the United States, and the United Kingdom. Of the 14 studies, 10 (71%) were RCTs, 2 (14%) were quasi-experimental studies, 1 (7%) was a controlled study, and 1 (7%) was a cohort study. Participants ranged from 28 to 433, with a total of 2339 participants across all included quantitative studies. Study populations included a general population of women who were pregnant (no psychological screening criteria), women with high anxiety and depression self-report scores, women with high fear of childbirth, women with insomnia, and preterm labor.

Study Rationale

In all the included studies, the authors stated that the effectiveness of cognitive-based or mindfulness approaches had been demonstrated in wider populations using traditional face-to-face sessions, with some evidence of effectiveness for internet-based CBT (I-CBT) approaches for anxiety and depression. Studies that focused on supporting women with fear of birth reported on the associations among anxiety, stress, and fear of birth. The authors hypothesized that CBT-based interventions would also lead to an improvement in women's fear of birth symptoms. Approximately 14% (2/14) of studies were focused on addressing birth outcomes. Urech et al [40] discussed the association between anxiety and preterm birth, suggesting psychological therapy as potentially helpful in improving birth outcomes. Shahsavan et al [41] reported an association between maternal requests for cesarean section and fear of birth, stating that addressing fear of birth symptoms may lead to a reduction in the prevalence of cesarean birth. Many of the included studies provided a rationale for developing remotely delivered interventions related to (1) flexibility, (2) cost-effectiveness, (3) potential reach, (4) accessibility, (5) availability, and (6) ubiquity. The flexibility of remotely delivered interventions was considered particularly relevant during pregnancy. The authors suggested that barriers to accessing traditional treatment, such as childcare issues, access to transportation, and limited time, could be overcome by providing interventions that women could access at the time of their choosing. Many studies have reported a lack of suitably trained therapists to meet the demand for psychological services, highlighting the accessibility of I-CBT as an advantage for increasing access to care for all women in all locations. Approximately 14% (2/14) of studies [17,42] discussed that remotely delivered interventions might provide women with greater anonymity and be more acceptable to women who may be reluctant to disclose their symptoms or seek support from health care professionals (HCPs) because of the stigma of mental health conditions in pregnancy, providing women with greater anonymity. The propensity of women who are pregnant for accessing web-based information and the ubiquity of smartphones, tablets, and computers was highlighted as further motivation for developing remotely delivered interventions for this population.

Comparators or CGs

Carissoli et al [43] conducted a pragmatic controlled trial, with the intervention group (IG) receiving a psychological well-being mobile app in addition to childbirth classes. The CG was randomized to receive only childbirth classes. Felder et al [44] recruited women with symptoms of insomnia. The CG was reported to receive standard care for insomnia, which the authors state may have included medication, psychotherapy, herbal supplements, counseling, and support groups. Urech et al [40] recruited women with diagnosed preterm labor for a cognitive behavioral stress management program. The CG received the same duration of web-based sessions and were asked to write short stories, listen to radio plays, answer specific questions, or perform quiz games. Other psychological, psychoeducation, and mindfulness intervention studies stated standard or usual antenatal care as the control condition [17,41,42,45-48,52]. The study by Rondung et al [49] compared I-CBT for women with fear of birth with standard care, which the authors described as including midwife-led fear of birth counseling, delivered face-to-face over 2 to 4 sessions. Kelman et al [50] compared internet-delivered compassionate mind training with I-CBT in a proof-of-concept RCT.

Intervention Components in the Included Studies

Interventions included psychological (CBT and problem solving), mindfulness, and educational designs. Many of the interventions adopted a multidimensional approach with cognitive behavioral techniques, relaxation, and educational intervention components. Most were delivered via web-based platforms where women could complete exercises and modules in their own time. Of the 13 interventions, 8 (62%) interventions included direct contact from trained therapists or coaches via feedback or answering questions, and 1 (8%) intervention was delivered via telephone [46]. Interventions lasted for 2 to 30 weeks (Multimedia Appendix 2). Participants were recruited by various methods: HCP referral during routine clinical appointments or self-recruited via advertisements in print and social media. Eligibility screening was conducted in 79% (11/14) of the studies, which included anxiety symptoms, depression, fear of birth, insomnia, or preterm labor (Multimedia Appendix 3). For studies reporting inclusion rates (as a percentage of the enrolled population), all 29% (4/14) of studies that used criteria based on fear of birth measures reported inclusion rates <10%. The 29% (4/14) of studies that used criteria based on anxiety or depression screening tools reported inclusion rates between 19% and 46%. Studies that did not apply psychological screening inclusion measures reported inclusion rates ranging from 76% to 90%.

Intervention Completion Rates Reported in the Included Studies

Shahsavan et al [41] reported an adherence rate of 93.72% across an 8-week I-CBT intervention. Adherence to the intervention program was measured as the number of log-ins and time spent on the designed software gathered via web - logging per unique user ID. Loughnan et al [53] and Yang et al [45] reported completion rates of 72% across all 3 sessions for an I-CBT intervention and 84% for a mindfulness intervention. The remaining 54% (7/13) of interventions that provided relevant

data reported falling attendance rates across the time frame of the intervention. Approximately 36% (5/14) of studies (I-CBT and problem-solving interventions) maintained completion rates of approximately ≥50% at week 5 to week 6 time points and included women with elevated symptoms of anxiety, depression, fear of birth, or insomnia [16,17,42,44,46]. Approximately 14% (2/14) of studies reported completion rates of <30% by week 3 on the intervention program: the study by Rondung et al [49], which involved I-CBT for women with high fear of birth scores, and the study by Krusche et al [47], which involved mindfulness for a general population of women who are pregnant. Fontein-Kuipers et al [48] reported the number of women who completed third trimester postintervention measures for an intervention delivering coping skills, resources, and personalized feedback (CG 56% and IG 64%). Across the 64% (9/14) of studies reporting full intervention completion data, all 56% (5/9) of studies that maintained completion rates >50% across the whole intervention reported significant improvement in anxiety scores between the IG and CG groups [16,41,44,46,53]. The 44% (4/9) of studies where completion fell <50% (at weeks 2-5) reported nonsignificant differences between the IG and CG groups [17,42,47,49].

Reported Acceptability and Satisfaction

Approximately 29% (4/14) of the studies assessed participant acceptability or satisfaction through feedback and satisfaction questionnaires; all received positive feedback, and women considered they had benefitted from participating. Loughnan et al [52] reported that after 1 to 2 sessions, most participants found the I-CBT intervention logical and considered it would be useful for their symptoms of anxiety and depression. Following completion of the intervention, most women were satisfied with the program, which they assessed as good quality, relatable, and useful in helping them manage their symptoms. Women could relate to the fictional characters experiencing anxiety and depression presented in an illustrated story. The authors reported that just over half of the participants preferred the web-based delivery of the program rather than other methods of intervention delivery. Forsell et al [17] reported that women who completed the I-CBT intervention felt satisfied, and most rated the treatment as helpful and important. Heller et al [42] reported that most women found the guided problem-solving intervention satisfying and would recommend the intervention to others. The website and feedback of coaches were rated as fairly good to excellent by most participants. The mindfulness intervention by Yang et al [45] was reported as beneficial by most of the participants who reported that it helped them feel relaxed and calm, become fully aware of fetal movements, and relieve discomfort. Some women reported that learning to maintain an accepting attitude was challenging, and daily mindfulness practice was onerous.

Reported Findings for Anxiety and Fear of Birth Scores in the Included Studies

Studies used various self-report outcome measures for anxiety and fear of birth symptoms, including the following:

- GAD 7-item (GAD-7) scale [18]
- Patient Health Questionnaire 9 item [54]
- Hospital Anxiety and Depression Scale—Anxiety subscale [55]
- Depression Anxiety Stress Scale [56]
- State–Trait Anxiety Inventory [57]
- Pregnancy-Related Anxiety Test [58]
- Pregnancy-Related Anxiety Questionnaire [59]
- Wijma Delivery Expectancy/Experience Questionnaire [60]
- Fear of Birth Scale [61]

Outcome scores as reported in the included papers are provided in Multimedia Appendix 3.

Meta-analysis of Anxiety Postintervention Scores

Studies used different anxiety measures to assess intervention outcomes (GAD-7, Hospital Anxiety and Depression Scale—Anxiety subscale, Pregnancy-Related Anxiety Test, Pregnancy-Related Anxiety Questionnaire, Patient Health Questionnaire-9 item, State–Trait Anxiety Inventory, and Depression Anxiety Stress Scale); therefore, SMDs were used as the summary statistic [33]. The results from 43% (6/14) of the studies that provided sufficient outcome data for postintervention anxiety scores were pooled and indicated statistical heterogeneity among the studies and clinical heterogeneity between the intervention type, duration, and the characteristics of participants (Figure 4). Subgroup analyses were conducted on studies of interventions with similar characteristics, such as I-CBT and facilitated interventions, which provided sufficient data and were assessed as having sufficient clinical and statistical homogeneity to perform a meta-analysis (Figures 5 and 6) [33].

For the 23% (3/13) I-CBT interventions that provided sufficient data to be included in the meta-analysis, a beneficial effect was observed in relation to the reduction of anxiety scores (SMD −0.49, 95% CI −0.75 to −0.22), with low statistical heterogeneity among the studies (I^2=0%; P=.87). For facilitated interventions, beneficial effects were observed in relation to the reduction of anxiety scores (SMD −0.48, 95% CI −0.75 to −0.22), although there was substantial statistical heterogeneity among the studies (I^2=65%; P=.02). However, the pooled number of participants was relatively small (n=228 and n=848, respectively), and studies were assessed to have an unclear risk of bias. Therefore, the results of the meta-analysis should be interpreted with caution (Table 1).

Figure 4. Studies reporting digitally delivered interventions for intervention and control groups (anxiety outcomes) [17,40,41,43,45,46,49,50,53].

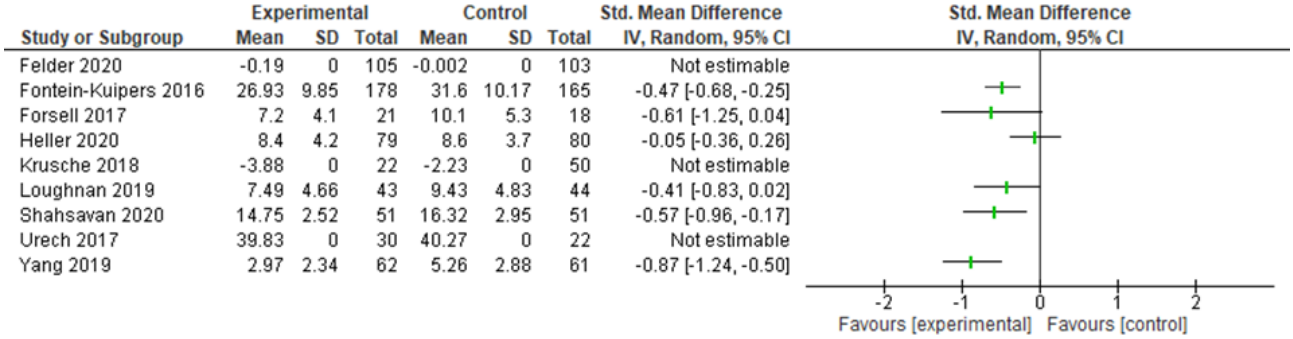


Figure 5. Studies reporting digitally delivered cognitive behavioral therapy interventions for intervention and control groups (anxiety outcomes) [17,40,41,45,52].

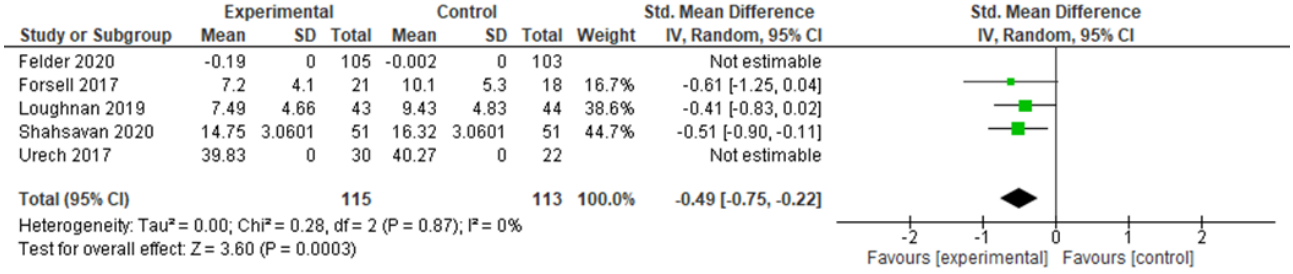


Figure 6. Studies reporting digitally delivered interventions with facilitator or therapist support for intervention and control groups (anxiety outcomes) [17,40,41,43,46,50].

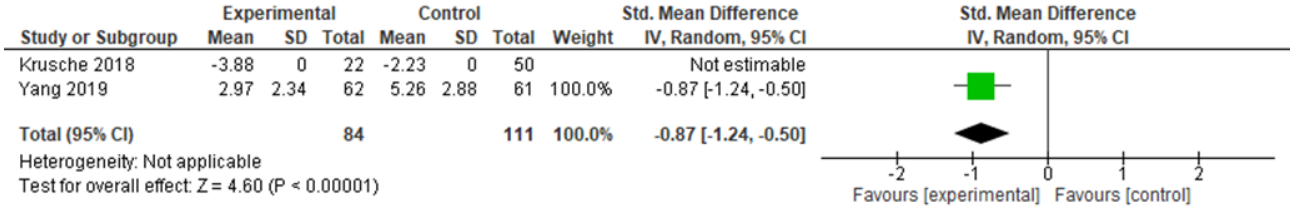


Table 1. GRADE[a] quality of evidence summary for I-CBT[b] and mindfulness RCTs[c] and quasi-experimental studies with anxiety score outcomes.

Interventions	Number of participants included in the analysis (N)	Quality of evidence (GRADE)	Intervention versus comparator mean difference (95% CI)
I-CBT interventions	228 from 2 RCTs and 1 quasi-experimental study	Very low because of small sample sizes, moderate risk of bias, differences in the population inclusion criteria, and intervention components	−0.49 (95% CI −0.75 to −0.22) improved anxiety scores in the intervention group
Facilitated interventions	818 from 3 RCTs and 2 quasi-experimental studies	Very low because of small sample sizes, moderate risk of bias, differences in the intervention types, population inclusion criteria, and intervention components	−0.48 (95% CI −0.75 to −0.22) improved anxiety scores in the intervention group

[a]GRADE: Grading of Recommendations Assessment, Development and Evaluation.

[b]I-CBT: internet-based cognitive behavioral therapy.

[c]RCT: randomized controlled trial.

Findings From the Qualitative Studies: Women's Engagement and Experiences of Interventions

A total of 3 qualitative studies were included in this review. Of the 3 studies, 2 (67%) included participants of experimental intervention studies [38,39], and 1 (33%) reported women's narratives from a web-based forum [51]. The study by Gui et al [51] was included in the review as the web-based forum was considered by the review team to fit with the eligibility criteria of an intervention, and the study provided useful and relevant data on women's concerns, anxiety, stress, and fear during pregnancy.

Baylis et al [39] interviewed women (n=19) who had participated in the I-CBT intervention to support women with a fear of childbirth [49]. The women were interviewed in the postnatal period. The study aimed to describe women's experiences of guided I-CBT for fear of childbirth and describe the content of that fear. The authors reported that women's fears were focused on losing control in labor, fear for their life, or

the health of the baby. I-CBT was reported to offer a supportive and flexible treatment, although most women would have preferred CBT to be delivered face-to-face. Remote delivery made it difficult for women to maintain their motivation. The I-CBT intervention was not tailored to women's specific situations or fears, and women could not always relate to the content within the sessions. However, women said it helped them to think in different ways about the imminent birth, understand their fear, work with their emotions, and cope with uncertainty. Learning the skills to *step outside* of themselves and observe what was going on helped women relieve their fears. The interactions with the psychologist through emails received positive comments and were described as supportive and encouraging. Individual contact made the treatment more personal, and the relationship with the psychologist was reported by some women as the most important part of the therapy. For some women, contact by email was preferable to meeting a stranger in an unfamiliar hospital setting and made it easier to disclose their fears. Other women would have preferred meeting the psychologist in person; this was thought to have resulted in women feeling more supported and motivated to complete the program and helped them create a warm and trusting relationship. Some women actively sought out alternative ways of addressing their fear of childbirth, including face-to-face midwife-led counseling, particularly when they needed to process previous traumatic birth experiences. Some women contacted a psychologist, attended antenatal classes, or attended antenatal yoga sessions.

Nieminen et al [38] conducted a thematic analysis of narratives completed by nulliparous women participating in an I-CBT intervention for fear of childbirth [16]. This study aimed to describe women's expectations of childbirth before and after I-CBT. The authors reported six main themes: fear, self-confidence, coping, the partner, the staff, and the baby. Themes were identified within three domains: my own role, the role of others, and attitude toward the baby. Women reported that, following I-CBT, their complete focus on fear, as well as anxiety and hopelessness, was replaced by an expectation that reflected increased confidence. Women described positive and more realistic expectations regarding themselves, their partners, and the staff that would look after them in labor. Before I-CBT, women provided very few comments about the health of the baby in labor and were more focused on their anxiety and on surviving painful labor and delivery, which some women described as a *nightmare*. Following I-CBT, women still reported uncertainty about the outcome of labor and birth; however, their attitude toward their fear had changed to become more manageable and less isolating. They developed self-confidence and strategies to cope with pain and seek support from partners and care providers.

The study by Gui et al [51] analyzed narratives from women who were pregnant and posting within a web-based community across the 3 trimesters of pregnancy. The study reported women's motivations for seeking web-based support, which surrounded (1) having limited access to care providers (delay in accessing advice); (2) wanting to access alternative sources of advice and information; (3) having little social support; (4) experiencing conflict between pregnancy information and their own experiences; and (5) seeking general advice, reassurance, and emotional support. The authors reported that across all 3 trimesters, women who were pregnant closely monitored their symptoms, fetal movements, body shape, and test results. Women reported emotional stress about maintaining their pregnancy, adjusting their lifestyle, managing medication for mental health conditions, and managing family relationships. Participation in web-based peer support provided a strategy for women to manage their stress and anxiety. Emotional stress remained a concern for women in the third trimester, which is mainly related to fear and anxiety about the upcoming birth, personal relationships with family members with regard to preparing for the birth, and extreme physical discomfort because of their progressing pregnancy. The authors stated that women's motivations for seeking web-based support were greater for women with particular symptoms who needed to access information immediately. The number of posts that contained statements about emotional stress was slightly larger than the number of posts that directly asked for emotional support, which the authors concluded was because of differences in the way women cope with stress (ie, some women actively seek social support, whereas others prefer to vent their feelings).

Synthesis of Quantitative and Qualitative Evidence

The evidence from the quantitative studies (14/16, 88%) and the qualitative studies (3/16, 18%) were synthesized to draw overall conclusions and provide insight to guide the design of future studies. The synthesis findings are presented in Textbox 2 and relate to the research questions.

Textbox 2. Synthesis of the qualitative and quantitative evidence.

Quantitative, qualitative, and synthesized evidence

How effective are remotely delivered interventions in reducing symptoms of anxiety?

- Quantitative evidence: There was limited evidence of effectiveness for cognitive behavioral therapy and facilitated interventions for women with symptoms of depression, anxiety, and fear of birth. Interventions may be more effective when interventions are targeted at women with psychological symptoms; interventions include facilitation and contact with therapists, health care professionals, or peer communities; completion rates are sustained above 50% across all sessions; and interventions include cognitive skills and relaxation or meditation.
- Qualitative evidence: There was limited evidence from 2 studies to suggest that internet-based cognitive behavioral therapy interventions can be beneficial in helping women develop coping strategies for dealing with uncertainties and accessing support from partners and health care professionals. Limitations of remotely delivered interventions include a lack of tailoring to specific fears and anxieties; a lack of motivation and encouragement to complete modules; an inability to form secure, trusting professional relationships; content is perceived as onerous.
- Synthesized findings: Overall, there is limited evidence to suggest that interventions are more effective when women are motivated to maintain regular participation in interventions. Strategies to maintain participation include the following:
  - Personalization: achieved by providing regular contact with therapists and health care professionals to discuss individual circumstances, symptoms, or concerns; consider animated therapists as a proxy for human therapist interactions
  - Relatable: designing and targeting interventions to women with particular symptoms; providing peer forums; presenting content as stories and vignettes
  - Skills and techniques: include components of relaxation and cognitive-based skills
  - Achievable: sufficient sessions to develop new skills without being too time consuming

What are women's views on remotely delivered interventions in reducing the symptoms of anxiety?

- Quantitative evidence: There is limited evidence to suggest that attrition rates were higher in studies that did not include therapist and health care professional or peer support components and were not targeted at women with symptoms of anxiety, depression, or fear of birth. Participants in 2 internet-based cognitive behavioral therapy interventions reported they were satisfied with the intervention and considered it helpful to manage their symptoms.
- Qualitative evidence: There was limited evidence from 2 studies to suggest that internet-based cognitive behavioral therapy interventions for fear of birth were valued by women if they offered flexibility and strategies for coping with their anxiety and fear and provided access to individual support from therapists. Women with fear of birth would have preferred individualized advice that was not readily available via remotely delivered therapy. Although some women felt more confident in disclosing their fears via remote messaging, many women would have preferred and actively sought alternative face-to-face treatment.
- Synthesized findings: Overall, there is very limited evidence to suggest that some women may prefer face-to-face therapy to help them stay motivated and receive individualized, tailored advice to manage their fears surrounding pregnancy and childbirth. However, for other women, remotely delivered interventions that provide some contact with a therapist, health care professional, or peer community may provide adequate human interaction to be of benefit. Remote interventions may also provide women with greater anonymity to help them feel more confident in disclosing their symptoms. There is limited evidence from the qualitative and quantitative evidence to suggest that women value interventions that resonate with their experiences. Women may be more motivated to complete interventions that are perceived as relevant to their needs and situations.

Discussion

Principal Findings

This paper presents a systematic review of remotely delivered interventions to improve symptoms of anxiety in women who are pregnant and presents a meta-analysis of the preliminary efficacy of these interventions on self-reported anxiety scores. Within the 16 included studies, numerous intervention designs were evaluated. Of the 16 studies, 10 (63%) evaluated I-CBT interventions targeting women with fear of birth, anxiety and depression, preterm birth, and insomnia; 2 (13%) reported mindfulness interventions for women who are pregnant and having symptoms of anxiety and depression; and the remaining 4 (25%) included psychoeducation for women with fear of birth, problem solving for women with anxiety and depression, psychological well-being, and a web-based forum for a general population of women who are pregnant. Remotely delivered interventions included in the meta-analysis achieved some beneficial effect in relation to the reduction of anxiety scores, although these findings need to be interpreted with caution as sample sizes were relatively small, and studies were assessed to have an unclear risk of bias.

The need to increase the reach and improve timely access to therapeutic and supportive treatment for women with psychological symptoms was reported as the main rationale for conducting the studies. In the United Kingdom, suicide is the second largest cause of maternal deaths, and women with a mental health diagnosis are overrepresented among women who die during pregnancy or the postnatal period [62]. Within the UK maternity services, low-level psychological therapies are recommended as the first-line treatment option for women with mild to moderate mental health conditions [63]. Improving Access to Psychological Therapy programs are available for many women who are pregnant in the United Kingdom, with recommended treatment interventions starting within 6 weeks of referral [64]. Women have reported positive experiences of Improving Access to Psychological Therapy support, although

barriers to accessing services have been identified, including reluctance to disclose difficulties because of stigma and fear of custody loss, as well as lack of clear information and support, with referrals from HCPs such as general practitioners, midwives, and health visitors [65]. This review has identified that for some women, remotely delivered interventions provide a sense of anonymity, which enables them to feel more confident in disclosing their symptoms. However, qualitative studies assessing women's feedback on remotely delivered interventions reported a lack of face-to-face contact with therapists and a lack of tailoring information and interventions to women's specific circumstances as barriers to the effectiveness of interventions. The quantitative data suggested that interventions were more effective in improving anxiety symptoms when interventions included individual contact from a therapist, HCP, or peer web-based forums.

Social learning and social comparison theory can be facilitated by providing social forums to foster a sense of connectivity and cooperation. Many women participate in web-based forums during pregnancy to discuss their worries and concerns and regularly access web-based information to monitor their well-being and assist in decision-making [51,66]. The inclusion of social support mechanisms as a component of remotely delivered interventions has been recommended for individuals with serious mental health concerns [67].

Therapeutic alliance has been reported as fundamental to the success of face-to-face psychological therapies and requires a strong bond between the care provider and client to foster shared understanding and collaboration on tasks and goals [67]. A systematic review of technology-based mental health interventions identified that users could experience therapeutic alliances with digital interventions if they are personalized and interactive, providing automated feedback that emulates reciprocal trusted relationships [68]. Numerous systematic reviews have identified that remotely delivered anxiety interventions are more effective when therapists provide support and guidance [69-71]. The impact of facilitator or therapist training has mainly been evaluated in digital interventions for symptoms of depression. Remotely delivered guided interventions were reported to be beneficial in improving symptoms of depression and found no association between the qualifications or level of facilitator training and outcomes [72-74]. In pregnancy, increasing midwives' and maternity care providers' awareness of remotely delivered interventions could assist signposting women to effective interventions. Remotely delivered interventions can be enhanced by midwives offering support and encouragement to improve intervention uptake and completion and providing pregnancy-specific information and advice tailored to women's particular circumstances. Midwives may need brief, additional training to support this role; however, in addition to supporting women's experience of the intervention, midwives' involvement would reflect contemporary policy drivers related to the continuity of carer (Better Births). A central tenet of the Midwifery Continuity of Carer models is the development of a collaborative relationship and provision of personalized care. Compared with conventional care, the provision of continuity of carers enables midwives to get to know women better, increase mutual trust, and facilitate access to specialist services.

Of the 16 included studies [44,53], 2 (13%) evaluated unguided I-CBT interventions for women with anxiety and depression and women with insomnia. Both studies reported an improvement in postintervention between-group anxiety symptoms for the IGs. Both studies included fictional characters in the form of a digital therapist [44] and fictional characters to narrate anxiety and depression experiences and help teach CBT skills. Rehm et al [75] reported that avatars were used as autonomous digital therapists to facilitate clinical interviews and assessments, provide psychoeducation, or signpost to other services. Digital therapists, not controlled by human clinicians, can be represented as a realistic-looking human avatar or as a 2D animated character. Participants have reported that they were willing and felt comfortable sharing information with therapist avatars [75], although there is very limited evidence, and further research is required in the perinatal context.

None of the included studies were solely targeted at improving symptoms of anxiety, which may reflect a greater focus on broader concepts of psychological well-being and respond to the reported comorbidity and associations between common mental health and psychosocial and physical health concerns in the perinatal period [76-78]. Although a multidimensional approach has been reported as an important factor in promoting psychological well-being in pregnancy [79], interventions targeting 1 condition may not be effective for the other comorbid conditions [80]; the underpinning theory of change needs to be defined for each condition before further testing the mechanisms of change.

The 31% (5/16) of studies that maintained completion rates ≥50% across the entire intervention reported significant improvement in anxiety scores between the IG and CG groups compared with studies where completion rates fell <50% (weeks 2-5). Of the 4 studies where completion fell <50%, 3 (75%) were developed for the general population of women who are pregnant and 1 (25%) for women with severe symptoms of fear of birth. Studies that maintained relatively high completion rates were targeted at women with symptoms of anxiety, depression, insomnia, or fear of birth and used established, validated screening tools with recommended cutoff scores to identify women with mild, moderate, and severe symptoms. A systematic review of antenatal interventions to reduce maternal distress also reported that interventions delivered to women with symptoms of distress were more effective than preventative interventions for women with little or no symptoms at baseline [81]. However, because of the heterogeneity of target populations included in the review, there is currently insufficient evidence to draw conclusions about the effectiveness of targeted interventions compared with universal interventions for women with symptoms of anxiety during pregnancy.

Overall, the content of interventions was well-documented, with most of the studies reporting the material content provided in each module. Of the 13 interventions, 10 (77%) included cognitive skills, and 69% (11/16) of studies included mindfulness or relaxation techniques. The amount of time taken to complete modules was not well-documented, and some

authors reported the average time women accessed websites and portals, whereas other authors reported the amount and types of information provided. Women reported that intervention content and homework exercises felt too onerous at times, and they found it difficult to remain motivated. In future studies, assessing and reporting the amount of time spent completing intervention modules would help inform optimal intervention content. Overall, the information about the ways in which interventions were developed and tailored to the population was limited, and many authors simply stated that interventions were amended from traditional CBT interventions to meet the needs of a pregnant population or to fit within the timescale of pregnancy. Conducting a needs assessment in the target population is an essential stage in intervention development to improve potential effectiveness [82]. O'Mahen et al [83] explored women's needs to inform the modification of CBT for perinatal depression. The authors reported that an increased focus on interpersonal strategies is required to help women seek out normalizing information that counters rigid beliefs about motherhood from other mothers. Considerations of women's ethnicity and socioeconomic status are also required to address particular worries and concerns and improve the relevance of the intervention content. Similar studies are required to inform the tailoring of interventions for women with symptoms of anxiety during pregnancy. Only 19% (3/16) of the included studies provided demographic data relating to the ethnicity of participants, and from these studies, most of the participants were White. The recent MBRRACE-UK: Saving Lives, Improving Mothers' Care report has highlighted inequalities in maternity care, with Black and Black British and Asian and Asian British babies up to twice as likely to be stillborn or die neonatally, suggesting that safety initiatives are failing to reach many women from higher risk ethnicities [84]. A low rate of participation among ethnic minority groups reduces the generalizability of mental health research findings, affects the development of effective interventions, and further widens health inequalities. Urgent attention is required to address inequalities in mental health provision for Black, Asian, and minority ethnic women, and researchers need to develop relevant, tailored interventions and recruitment strategies that reflect the diversity of the population.

Limitations

This review adopted a broad approach, including different types of interventions for different populations of women who are pregnant. This limited the utility of the study findings as different measurement instruments were used by the authors, and there was insufficient data to calculate outcome scores across all included studies. As a meta-analysis of postintervention anxiety scores was only achievable for a small subgroup of studies, the aim of the study to assess the effectiveness of interventions was only partially achieved. However, low-powered analysis based on a small number of studies can provide useful insights by exploring interesting relationships that may emerge and highlighting deficiencies in a topic that requires further attention [85]. The overall quality of the included studies was rated as moderate; the small sample sizes and heterogeneity of interventions and study populations resulted in overall low quality of evidence for the quantitative studies. There were insufficient qualitative studies to conduct a narrative synthesis, which further limited the qualitative findings and the synthesis of qualitative and quantitative evidence. We did not locate any studies that focused solely on anxiety symptoms in pregnancy, and most quantitative studies were not powered to detect significant changes in anxiety scores. Studies not published in English were excluded from this review.

Conclusions

The introduction of remotely delivered interventions has the potential to improve symptoms of anxiety in women who are pregnant. The results of the review are limited and need to be interpreted with caution, as the findings of the review were predominantly based on small sample sizes with heterogeneity between intervention designs, delivery, and sample populations. The synthesized findings highlighted components of interventions that may improve the effectiveness and acceptability of remotely delivered interventions. Most women valued individual web-based contact from a therapist or HCP to maintain motivation and access individualized information. There was some evidence of effectiveness for interventions that provided some form of facilitation and access to peer support. Overall, there was limited evidence to suggest that interventions are more effective when women are motivated to maintain regular participation. Interventions targeting women with psychological symptoms were more likely to maintain participation across the intervention time frame and report improvements in anxiety scores. Although there was some limited evidence of the benefits of remotely delivered interventions described as CBT or mindfulness, most studies included a multicomponent approach and provided cognitive and mind–body content. Most women reported satisfaction with the interventions and provided positive feedback. Some women may prefer and actively seek face-to-face therapeutic interventions. However, for some women, remotely delivered interventions provide women with greater anonymity to help them feel more confident in disclosing their symptoms. Interventions provided a timely and flexible approach and provided women with strategies for coping with their symptoms. Future research is required to identify ways that interventions can be tailored to meet the particular needs of diverse populations of women who are pregnant to improve their relevance and ensure equitable access. Researchers need to consider the structures of maternity care in which interventions are implemented to assist in signposting women and need to maximize the potential for maternity care professionals to provide encouragement, support, and motivation, enhancing the digital therapeutic approach.

Authors' Contributions

KE and SC reviewed and assessed the papers for submission. All authors contributed to the study design, analysis, and review of the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Summary of the intervention studies included in the review: TIDieR (Template for Intervention Description and Replication) checklist (Hoffmann et al [37]).
[DOCX File , 65 KB - jmir_v24i2e28093_app1.docx]

Multimedia Appendix 2
Intervention components in the included studies.
[DOCX File , 69 KB - jmir_v24i2e28093_app2.docx]

Multimedia Appendix 3
Recruitment methods, inclusion and completion rates, and outcome measures reported in the included studies.
[DOCX File , 81 KB - jmir_v24i2e28093_app3.docx]

References

1. Figueiredo B, Conde A. Anxiety and depression in women and men from early pregnancy to 3-months postpartum. Arch Womens Ment Health 2011;14(3):247-255. [doi: 10.1007/s00737-011-0217-3] [Medline: 21479759]
2. Martini J, Wittich J, Petzoldt J, Winkel S, Einsle F, Siegert J, et al. Maternal anxiety disorders prior to conception, psychopathology during pregnancy and early infants' development: a prospective-longitudinal study. Arch Womens Ment Health 2013;16(6):549-560. [doi: 10.1007/s00737-013-0376-5] [Medline: 24057868]
3. Dennis CL, Falah-Hassani K, Shiri R. Prevalence of antenatal and postnatal anxiety: systematic review and meta-analysis. Br J Psychiatry 2017;210(5):315-323. [doi: 10.1192/bjp.bp.116.187179] [Medline: 28302701]
4. Fisher J, Cabral de Mello M, Patel V, Rahman A, Tran T, Holton S, et al. Prevalence and determinants of common perinatal mental disorders in women in low- and lower-middle-income countries: a systematic review. Bull World Health Organ 2012;90(2):139G-149G [FREE Full text] [doi: 10.2471/BLT.11.091850] [Medline: 22423165]
5. Fawcett EJ, Fairbrother N, Cox ML, White IR, Fawcett JM. The prevalence of anxiety disorders during pregnancy and the postpartum period: a multivariate Bayesian meta-analysis. J Clin Psychiatry 2019;80(4):18r12527 [FREE Full text] [doi: 10.4088/JCP.18r12527] [Medline: 31347796]
6. Leach LS, Poyser C, Fairweather‐Schmidt K. Maternal perinatal anxiety: a review of prevalence and correlates. Clin Psychol 2017;21(1):4-19. [doi: 10.1111/cp.12058]
7. Fontein-Kuipers Y, Ausems M, Budé L, Van Limbeek E, De Vries R, Nieuwenhuijze M. Factors influencing maternal distress among Dutch women with a healthy pregnancy. Women Birth 2015;28(3):e36-e43. [doi: 10.1016/j.wombi.2015.02.002] [Medline: 25757732]
8. Staneva A, Bogossian F, Pritchard M, Wittkowski A. The effects of maternal depression, anxiety, and perceived stress during pregnancy on preterm birth: a systematic review. Women Birth 2015;28(3):179-193. [doi: 10.1016/j.wombi.2015.02.003] [Medline: 25765470]
9. Grigoriadis S, Graves L, Peer M, Mamisashvili L, Tomlinson G, Vigod SN, et al. A systematic review and meta-analysis of the effects of antenatal anxiety on postpartum outcomes. Arch Womens Ment Health 2019;22(5):543-556. [doi: 10.1007/s00737-018-0930-2] [Medline: 30523416]
10. Generalised anxiety disorder and panic disorder in adults: management. National Institute for Health and Care Excellence. 2019. URL: https://www.nice.org.uk/guidance/cg113 [accessed 2022-01-25]
11. Howard LM, Khalifeh H. Perinatal mental health: a review of progress and challenges. World Psychiatry 2020;19(3):313-327 [FREE Full text] [doi: 10.1002/wps.20769] [Medline: 32931106]
12. Vigod SN, Dennis CL. Advances in virtual care for perinatal mental disorders. World Psychiatry 2020;19(3):328-329 [FREE Full text] [doi: 10.1002/wps.20775] [Medline: 32931113]
13. Silver L. Smartphone ownership is growing rapidly around the world, but not always equally. Pew Research Center. Washington, D.C: Pew Research Center; 2019. URL: https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/ [accessed 2022-01-25]
14. Hussain-Shamsy N, Shah A, Vigod SN, Zaheer J, Seto E. Mobile health for perinatal depression and anxiety: scoping review. J Med Internet Res 2020;22(4):e17011 [FREE Full text] [doi: 10.2196/17011] [Medline: 32281939]

15. Loughnan SA, Joubert AE, Grierson A, Andrews G, Newby JM. Internet-delivered psychological interventions for clinical anxiety and depression in perinatal women: a systematic review and meta-analysis. Arch Womens Ment Health 2019;22(6):737-750. [doi: 10.1007/s00737-019-00961-9] [Medline: 31101993]
16. Nieminen K, Andersson G, Wijma B, Ryding EL, Wijma K. Treatment of nulliparous women with severe fear of childbirth via the Internet: a feasibility study. J Psychosom Obstet Gynaecol 2016;37(2):37-43. [doi: 10.3109/0167482X.2016.1140143] [Medline: 26918752]
17. Forsell E, Bendix M, Holländare F, Szymanska von Schultz B, Nasiell J, Blomdahl-Wetterholm M, et al. Internet delivered cognitive behavior therapy for antenatal depression: a randomised controlled trial. J Affect Disord 2017;221:56-64 [FREE Full text] [doi: 10.1016/j.jad.2017.06.013] [Medline: 28628768]
18. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
19. Wu H, Sun W, Huang X, Yu S, Wang H, Bi X, et al. Online antenatal care during the COVID-19 pandemic: opportunities and challenges. J Med Internet Res 2020;22(7):e19916 [FREE Full text] [doi: 10.2196/19916] [Medline: 32658860]
20. Peahl AF, Smith RD, Moniz MH. Prenatal care redesign: creating flexible maternity care models through virtual care. Am J Obstet Gynecol 2020;223(3):389.e1-389.10 [FREE Full text] [doi: 10.1016/j.ajog.2020.05.029] [Medline: 32425200]
21. Jardine J, Relph S, Magee LA, von Dadelszen P, Morris E, Ross-Davie M, et al. Maternity services in the UK during the coronavirus disease 2019 pandemic: a national survey of modifications to standard care. BJOG 2021;128(5):880-889. [doi: 10.1111/1471-0528.16547] [Medline: 32992408]
22. Rhodes A, Kheireddine S, Smith AD. Experiences, attitudes, and needs of users of a pregnancy and parenting app (Baby Buddy) during the COVID-19 pandemic: mixed methods study. JMIR Mhealth Uhealth 2020;8(12):e23157 [FREE Full text] [doi: 10.2196/23157] [Medline: 33264100]
23. Chivers BR, Garad RM, Boyle JA, Skouteris H, Teede HJ, Harrison CL. Perinatal distress during COVID-19: thematic analysis of an online parenting forum. J Med Internet Res 2020;22(9):e22002 [FREE Full text] [doi: 10.2196/22002] [Medline: 32857707]
24. Aromataris E, Munn Z. JBI manual for evidence synthesis. Adelaide: Joanna Briggs Institute; 2020.
25. Pluye P, Hong QN. Combining the power of stories and the power of numbers: mixed methods research and mixed studies reviews. Annu Rev Public Health 2014;35:29-45. [doi: 10.1146/annurev-publhealth-032013-182440] [Medline: 24188053]
26. Lizarondo L, Stern C, Carrier J, Godfrey C, Rieger K, Salmond S, et al. Chapter 8: Mixed methods systematic reviews. In: Aromataris E, Munn Z, editors. JBI manual for evidence synthesis. Adelaide: Joanna Briggs Institute; 2020.
27. Sandelowski M, Voils CI, Barroso J. Defining and designing mixed research synthesis studies. Res Sch 2006;13(1):29 [FREE Full text] [Medline: 20098638]
28. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
29. EPOC resources for review authors. Cochrane Effective Practice and Organisation of Care. 2017. URL: https://epoc.cochrane.org/resources/epoc-resources-review-authors [accessed 2022-01-25]
30. CASP qualitative checklist. Critical appraisal skills programme. 2018. URL: https://casp-uk.net/casp-tools-checklists/ [accessed 2021-05-15]
31. Quality assessment tool for before-after (pre-post) studies with no control group. National Heart, Lung, and Blood Institute. URL: https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools [accessed 2021-05-15]
32. Checklist for quasi-experimental studies. Joanna Briggs Institute. 2017. URL: https://jbi.global/critical-appraisal-tools [accessed 2022-01-25]
33. Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane handbook for systematic reviews of interventions version 6.2 (updated February 2021). London: Cochrane; 2021.
34. Schünemann HJ, Higgins JPT, Vist GE, Glasziou P, Akl EA, Skoetz N, et al. Chapter 14: Completing 'Summary of findings' tables and grading the certainty of the evidence. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021). London: Cochrane; 2021. URL: http://www.training.cochrane.org/handbook
35. Lewin S, Booth A, Glenton C, Munthe-Kaas H, Rashidian A, Wainwright M, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings: introduction to the series. Implement Sci 2018;13(Suppl 1):2 [FREE Full text] [doi: 10.1186/s13012-017-0688-3] [Medline: 29384079]
36. Sandelowski M, Voils CI, Leeman J, Crandell JL. Mapping the mixed methods-mixed research synthesis terrain. J Mix Methods Res 2012;6(4):317-331 [FREE Full text] [doi: 10.1177/1558689811427913] [Medline: 23066379]
37. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014;348:g1687. [doi: 10.1136/bmj.g1687] [Medline: 24609605]
38. Nieminen K, Malmquist A, Wijma B, Ryding EL, Andersson G, Wijma K. Nulliparous pregnant women's narratives of imminent childbirth before and after internet-based cognitive behavioural therapy for severe fear of childbirth: a qualitative study. BJOG 2015;122(9):1259-1265. [doi: 10.1111/1471-0528.13358] [Medline: 25817045]

39. Baylis R, Ekdahl J, Haines H, Rubertsson C. Women's experiences of Internet-delivered Cognitive Behaviour Therapy (iCBT) for fear of birth. Women Birth 2020;33(3):e227-e233. [doi: 10.1016/j.wombi.2019.05.006] [Medline: 31160244]
40. Urech C, Scherer S, Emmenegger M, Gaab J, Tschudin S, Hoesli I, et al. Efficacy of an internet-based cognitive behavioral stress management training in women with idiopathic preterm labor: a randomized controlled intervention study. J Psychosom Res 2017;103:140-146. [doi: 10.1016/j.jpsychores.2017.10.014] [Medline: 29167041]
41. Shahsavan F, Akbari N, Gharraee B, Abolghasemi J, Khedmat L. The effect of internet-based guided self-help cognitive-behavioral therapies on Iranian women's psychological symptoms and preferred method of childbirth. Perspect Psychiatr Care 2021;57(1):138-147. [doi: 10.1111/ppc.12535] [Medline: 32567051]
42. Heller HM, Hoogendoorn AW, Honig A, Broekman BF, van Straten A. The effectiveness of a guided internet-based tool for the treatment of depression and anxiety in pregnancy (MamaKits Online): randomized controlled trial. J Med Internet Res 2020;22(3):e15172 [FREE Full text] [doi: 10.2196/15172] [Medline: 32202505]
43. Carissoli C, Villani D, Gasparri D, Riva G. Enhancing psychological wellbeing of women approaching the childbirth: a controlled study with a mobile application. Annu Rev CyberTherapy Telemed 2017;15:45-50.
44. Felder JN, Epel ES, Neuhaus J, Krystal AD, Prather AA. Efficacy of Digital Cognitive Behavioral Therapy for the treatment of insomnia symptoms among pregnant women: a randomized clinical trial. JAMA Psychiatry 2020;77(5):484-492 [FREE Full text] [doi: 10.1001/jamapsychiatry.2019.4491] [Medline: 31968068]
45. Yang M, Jia G, Sun S, Ye C, Zhang R, Yu X. Effects of an online mindfulness intervention focusing on attention monitoring and acceptance in pregnant women: a randomized controlled trial. J Midwifery Womens Health 2019;64(1):68-77. [doi: 10.1111/jmwh.12944] [Medline: 30695166]
46. Toohill J, Fenwick J, Gamble J, Creedy DK, Buist A, Turkstra E, et al. A randomized controlled trial of a psycho-education intervention by midwives in reducing childbirth fear in pregnant women. Birth 2014;41(4):384-394 [FREE Full text] [doi: 10.1111/birt.12136] [Medline: 25303111]
47. Krusche A, Dymond M, Murphy SE, Crane C. Mindfulness for pregnancy: a randomised controlled study of online mindfulness during pregnancy. Midwifery 2018;65:51-57. [doi: 10.1016/j.midw.2018.07.005] [Medline: 30099285]
48. Fontein-Kuipers YJ, Ausems M, de Vries R, Nieuwenhuijze MJ. The effect of Wazzup Mama?! An antenatal intervention to prevent or reduce maternal distress in pregnancy. Arch Womens Ment Health 2016;19(5):779-788. [doi: 10.1007/s00737-016-0614-8] [Medline: 26965708]
49. Rondung E, Ternström E, Hildingsson I, Haines HM, Sundin Ö, Ekdahl J, et al. Comparing Internet-based Cognitive Behavioral Therapy with standard care for women with fear of birth: randomized controlled trial. JMIR Ment Health 2018;5(3):e10420 [FREE Full text] [doi: 10.2196/10420] [Medline: 30097422]
50. Kelman AR, Evare BS, Barrera AZ, Muñoz RF, Gilbert P. A proof-of-concept pilot randomized comparative trial of brief Internet-based compassionate mind training and cognitive-behavioral therapy for perinatal and intending to become pregnant women. Clin Psychol Psychother 2018;25(4):608-619. [doi: 10.1002/cpp.2185] [Medline: 29473698]
51. Gui X, Chen Y, Kou Y, Pine K, Chen Y. Investigating support seeking from peers for pregnancy in online health communities. Proc ACM Hum-Comput Interact 2017;1(CSCW):50.1-5019. [doi: 10.1145/3134685]
52. Loughnan SA, Newby JM, Haskelberg H, Mahoney A, Kladnitski N, Smith J, et al. Internet-based cognitive behavioural therapy (iCBT) for perinatal anxiety and depression versus treatment as usual: study protocol for two randomised controlled trials. Trials 2018;19(1):56 [FREE Full text] [doi: 10.1186/s13063-017-2422-5] [Medline: 29357918]
53. Loughnan SA, Sie A, Hobbs MJ, Joubert AE, Smith J, Haskelberg H, et al. A randomized controlled trial of 'MUMentum Pregnancy': Internet-delivered cognitive behavioral therapy program for antenatal anxiety and depression. J Affect Disord 2019;243:381-390. [doi: 10.1016/j.jad.2018.09.057] [Medline: 30266030]
54. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
55. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983;67(6):361-370. [doi: 10.1111/j.1600-0447.1983.tb09716.x] [Medline: 6880820]
56. Lovibond SH, Lovibond PF, Psychology Foundation of Australia. Manual for the depression anxiety stress scales. 2nd edition. Sydney: Psychology Foundation of Australia; 1995.
57. Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA. Manual for the state-trait anxiety inventory. Palo Alto: Consulting Psychologists Press; 1983.
58. Rini CK, Dunkel-Schetter C, Wadhwa PD, Sandman CA. Psychological adaptation and birth outcomes: the role of personal resources, stress, and sociocultural context in pregnancy. Health Psychology 1999;18(4):333-345. [doi: 10.1037//0278-6133.18.4.333] [Medline: 10431934]
59. Van den Bergh BR. The influence of maternal emotions during pregnancy on fetal and neonatal behavior. J Prenat Perinat Psychol Health 1990;5(2):119-130.
60. Wijma K, Wijma B, Zar M. Psychometric aspects of the W-DEQ; a new questionnaire for the measurement of fear of childbirth. J Psychosom Obstet Gynaecol 1998;19(2):84-97. [doi: 10.3109/01674829809048501] [Medline: 9638601]
61. Haines H, Pallant JF, Karlström A, Hildingsson I. Cross-cultural comparison of levels of childbirth-related fear in an Australian and Swedish sample. Midwifery 2011;27(4):560-567. [doi: 10.1016/j.midw.2010.05.004] [Medline: 20598787]

62. Knight M, Bunch K, Tuffnell D, Shakespeare J, Kotnis R, Kenyon S, et al. Saving lives, improving mothers' care: lessons learned to inform maternity care from the UK and Ireland confidential enquiries into maternal deaths and morbidity 2015-17. MBRRACE-UK. 2019. URL: https://www.npeu.ox.ac.uk/assets/downloads/mbrrace-uk/reports/MBRRACE-UK%20Maternal%20Report%202019%20-%20WEB%20VERSION.pdf [accessed 2022-01-25]
63. Antenatal and postnatal mental health: clinical management and service guidance. National Institute for Heath and Care Excellence. 2014. URL: https://www.nice.org.uk/guidance/cg192 [accessed 2022-01-25]
64. The perinatal mental health care pathways: full implementation guidance. National Collaborating Centre for Mental Health. 2018. URL: https://www.rcpsych.ac.uk/docs/default-source/improving-care/nccmh/perinatal/nccmh-the-perinatal-mental-health-care-pathways-full-implementation-guidance.pdf?sfvrsn=73c19277_2 [accessed 2022-01-25]
65. Millett L, Taylor BL, Howard LM, Bick D, Stanley N, Johnson S. Experiences of improving access to psychological therapy services for perinatal mental health difficulties: a qualitative study of women's and therapists' views. Behav Cogn Psychother 2018;46(4):421-436. [doi: 10.1017/S1352465817000650] [Medline: 29081328]
66. Wexler A, Davoudi A, Weissenbacher D, Choi R, O'Connor K, Cummings H, et al. Pregnancy and health in the age of the Internet: a content analysis of online "birth club" forums. PLoS One 2020;15(4):e0230947 [FREE Full text] [doi: 10.1371/journal.pone.0230947] [Medline: 32287266]
67. Tremain H, McEnery C, Fletcher K, Murray G. The therapeutic alliance in digital mental health interventions for serious mental illnesses: narrative review. JMIR Ment Health 2020;7(8):e17204 [FREE Full text] [doi: 10.2196/17204] [Medline: 32763881]
68. Hillier L. Exploring the nature of the therapeutic alliance in technology-based interventions for mental health problems. Lancaster University. Lancaster: Lancaster University; 2018. URL: https://www.research.lancs.ac.uk/portal/en/publications/exploring-the-nature-of-the-therapeutic-alliance-in-technologybased-interventions-for-mental-health-problems(406f0a3e-31d2-4295-81fd-c9d7e74659fe)/export.html [accessed 2022-01-25]
69. Saddichha S, Al-Desouki M, Lamia A, Linden IA, Krausz M. Online interventions for depression and anxiety - a systematic review. Health Psychol Behav Med 2014;2(1):841-881 [FREE Full text] [doi: 10.1080/21642850.2014.945934] [Medline: 25750823]
70. Palmqvist B, Carlbring P, Andersson G. Internet-delivered treatments with or without therapist input: does the therapist factor have implications for efficacy and cost? Expert Rev Pharmacoecon Outcomes Res 2007;7(3):291-297. [doi: 10.1586/14737167.7.3.291] [Medline: 20528315]
71. Spek V, Cuijpers P, Nyklícek I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. Psychol Med 2007;37(3):319-328. [doi: 10.1017/S0033291706008944] [Medline: 17112400]
72. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on Internet-based mental health interventions — a systematic review. Internet Interv 2014;1(4):205-215. [doi: 10.1016/j.invent.2014.08.003]
73. Titov N, Andrews G, Davies M, McIntyre K, Robinson E, Solley K. Internet treatment for depression: a randomized controlled trial comparing clinician vs. technician assistance. PLoS One 2010;5(6):e10939 [FREE Full text] [doi: 10.1371/journal.pone.0010939] [Medline: 20544030]
74. Gellatly J, Bower P, Hennessy S, Richards D, Gilbody S, Lovell K. What makes self-help interventions effective in the management of depressive symptoms? Meta-analysis and meta-regression. Psychol Med 2007;37(9):1217-1228. [doi: 10.1017/S0033291707000062] [Medline: 17306044]
75. Rehm IC, Foenander E, Wallace K, Abbott JM, Kyrios M, Thomas N. What role can avatars play in e-mental health interventions? Exploring new models of client-therapist interaction. Front Psychiatry 2016;7:186 [FREE Full text] [doi: 10.3389/fpsyt.2016.00186] [Medline: 27917128]
76. Biaggi A, Conroy S, Pawlby S, Pariante CM. Identifying the women at risk of antenatal anxiety and depression: a systematic review. J Affect Disord 2016;191:62-77 [FREE Full text] [doi: 10.1016/j.jad.2015.11.014] [Medline: 26650969]
77. Lavender TJ, Ebert L, Jones D. An evaluation of perinatal mental health interventions: an integrative literature review. Women Birth 2016;29(5):399-406. [doi: 10.1016/j.wombi.2016.04.004] [Medline: 27118000]
78. Wadephul F, Glover L, Jomeen J. Conceptualising women's perinatal well-being: a systematic review of theoretical discussions. Midwifery 2020;81:102598. [doi: 10.1016/j.midw.2019.102598] [Medline: 31835103]
79. Jomeen J. The importance of assessing psychological status during pregnancy, childbirth and the postnatal period as a multidimensional construct: a literature review. Clin Eff Nurs 2004;8(3-4):143-155 [FREE Full text] [doi: 10.1016/j.cein.2005.02.001]
80. Garber J, Weersing VR. Comorbidity of anxiety and depression in youth: implications for treatment and prevention. Clin Psychol (New York) 2010;17(4):293-306 [FREE Full text] [doi: 10.1111/j.1468-2850.2010.01221.x] [Medline: 21499544]
81. Fontein-Kuipers YJ, Nieuwenhuijze MJ, Ausems M, Budé L, de Vries R. Antenatal interventions to reduce maternal distress: a systematic review and meta-analysis of randomised trials. BJOG 2014;121(4):389-397 [FREE Full text] [doi: 10.1111/1471-0528.12500] [Medline: 24397691]
82. Bartholomew Eldredge LK, Parcel GS, Kok G, Gottlieb NH, Schaalma H, Markham CM. Planning health promotion programs: an intervention mapping approach. 2nd edition. New York: Wiley; 2011.

83. O'Mahen H, Fedock G, Henshaw E, Himle JA, Forman J, Flynn HA. Modifying CBT for perinatal depression: what do women want?: a qualitative study. Cogn Behav Pract 2012;19(2):359-371. [doi: 10.1016/j.cbpra.2011.05.005]
84. Draper ES, Gallimore ID, Smith LK, Fenton AC, Kurinczuk JJ, Smith PW, MBRRACE-UK collabloration. MBRRACE-UK perinatal mortality surveillance report: UK perinatal deaths for births from January to December 2018. MBRRACE-UK. 2020. URL: https://www.npeu.ox.ac.uk/assets/downloads/mbrrace-uk/reports/perinatal-surveillance-report-2018/MBRRACE-UK_Perinatal_Surveillance_Report_2018_-_final_v3.pdf [accessed 2022-01-25]
85. Greco T, Zangrillo A, Biondi-Zoccai G, Landoni G. Meta-analysis: pitfalls and hints. Heart Lung Vessel 2013;5(4):219-225 [FREE Full text] [Medline: 24364016]

Abbreviations

CBT: cognitive behavioral therapyCG: control group
GAD: Generalized Anxiety Disorder
GAD-7: Generalized Anxiety Disorder 7 items
GRADE: Grading of Recommendations Assessment, Development, and Evaluation
GRADE-CERQual: GRADE Confidence in Evidence from Reviews of Qualitative research
HCP: health care professional
I-CBT: internet-based cognitive behavioral therapy
IG: intervention group
JBI: Joanna Briggs Institute
PRISMA: Preferred Reporting Item for Systematic Reviews and Meta-Analyses
PROSPERO: The International Prospective Register of Systematic Reviews
RCT: randomized controlled trial
SMD: standardized mean difference
TIDieR: Template for Intervention Description and Replication

Edited by G Eysenbach; submitted 20.02.21; peer-reviewed by H Heller, N Hussain-Shamsy; comments to author 10.05.21; revised version received 03.06.21; accepted 19.11.21; published 15.02.22.

Please cite as:
Evans K, Rennick-Egglestone S, Cox S, Kuipers Y, Spiby H
Remotely Delivered Interventions to Support Women With Symptoms of Anxiety in Pregnancy: Mixed Methods Systematic Review and Meta-analysis
J Med Internet Res 2022;24(2):e28093
URL: https://www.jmir.org/2022/2/e28093
doi:10.2196/28093
PMID:35166688

Review

Online Health Information Seeking Behaviors Among Older Adults: Systematic Scoping Review

Yuxiang Chris Zhao[1], PhD; Mengyuan Zhao[1], MSc; Shijie Song[2], PhD

[1]School of Economics and Management, Nanjing University of Science and Technology, Nanjing, China

[2]Business School, Hohai University, Nanjing, China

Corresponding Author:
Shijie Song, PhD
Business School
Hohai University
Fo-Cheng West Rd 8
Nanjing, 211000
China
Phone: 86 15951973800
Email: ssong@hhu.edu.cn

Abstract

Background: With the world's population aging, more health-conscious older adults are seeking health information to make better-informed health decisions. The rapid growth of the internet has empowered older adults to access web-based health information sources. However, research explicitly exploring older adults' online health information seeking (OHIS) behavior is still underway.

Objective: This systematic scoping review aims to understand older adults' OHIS and answer four research questions: (1) What types of health information do older adults seek and where do they seek health information on the internet? (2) What are the factors that influence older adults' OHIS? (3) What are the barriers to older adults' OHIS? (4) How can we intervene and support older adults' OHIS?

Methods: A comprehensive literature search was performed in November 2020, involving the following academic databases: Web of Science; Cochrane Library database; PubMed; MEDLINE; CINAHL Plus; APA PsycINFO; Library and Information Science Source; Library, Information Science and Technology Abstracts; Psychology and Behavioral Sciences Collection; Communication & Mass Media Complete; ABI/INFORM; and ACM Digital Library. The initial search identified 8047 publications through database search strategies. After the removal of duplicates, a data set consisting of 5949 publications was obtained for screening. Among these, 75 articles met the inclusion criteria. Qualitative content analysis was performed to identify themes related to the research questions.

Results: The results suggest that older adults seek 10 types of health information from 6 types of internet-based information sources and that 2 main categories of influencing factors, individual-related and source-related, impact older adults' OHIS. Moreover, the results reveal that in their OHIS, older adults confront 3 types of barriers, namely individual, social, and those related to information and communication technologies. Some intervention programs based on educational training workshops have been created to intervene and support older adults' OHIS.

Conclusions: Although OHIS has become increasingly common among older adults, the review reveals that older adults' OHIS behavior is not adequately investigated. The findings suggest that more studies are needed to understand older adults' OHIS behaviors and better support their medical and health decisions in OHIS. Based on the results, the review proposes multiple objectives for future studies, including (1) more investigations on the OHIS behavior of older adults above 85 years; (2) conducting more longitudinal, action research, and mixed methods studies; (3) elaboration of the mobile context and cross-platform scenario of older adults' OHIS; (4) facilitating older adults' OHIS by explicating technology affordance; and (5) promoting and measuring the performance of OHIS interventions for older adults.

(J Med Internet Res 2022;24(2):e34790) doi:10.2196/34790

KEYWORDS

older adults; online health information seeking; health information behavior; aging technology; systematic scoping review

Introduction

During the past decade, the rapid development of information and communication technologies (ICTs) has increased laypeople's access to health information sources and is constantly reshaping their health information–seeking behaviors [1]. Online health information seeking (OHIS) serves multiple purposes, such as understanding disease symptoms, assessing disease risks, finding treatment choices, managing chronic conditions, and preparing for patient-doctor communication [2]. Studies have revealed that OHIS has become one of the most common everyday life experiences across the entire lifespan [3].

In recent decades, the aging of the world population has led to significant demographic transitions that have never occurred before in human history. Societies with large aging populations face great challenges to their health care sectors with respect to an increasing prevalence of chronic conditions among older adults and a sharply rising demand for health care resources. As older adults are more likely to experience illness and chronic conditions than younger people, they have a greater need for health information [4]. With the world population aging, increasing numbers of health-conscious older adults are seeking health information to make better-informed health decisions [5]. Many hopes are placed on ICTs to empower the aging population, promote public health, and alleviate the burden of health care systems. However, there is some skepticism regarding whether older adults really benefit from current technological advancements [6]. Although some studies have found that the adoption and use of ICTs to address health concerns have remained at a relatively low rate among older adults [7], other studies suggest that older adults are increasingly engaged in internet surfing [8]. These mixed results suggest that the OHIS behavior of older adults is still insufficiently investigated.

Despite scattered empirical studies on the topic, few scoping or systematic reviews have directly addressed the OHIS behaviors of older adults and synthesized this body of knowledge. Chang and Huang [9] recently reviewed antecedents that predict general consumers' OHIS behaviors (ie, health status, self-efficacy, health literacy, availability, credibility, emotional responses, and subject norms). Although the review found that age is a significant moderator of the correlations between the antecedents and OHIS, it provided few details on older adults' health information behaviors. Hunsaker and Hargittai [8] synthesized quantitative literature on general internet use among older adults. Although their review addressed the relationship between older adults' health and internet use, OHIS was neither specified nor teased out from the general internet use behaviors. Therefore, the type of health information sought by the participating older adults and the factors that influenced older adults' OHIS reported in the literature are unclear. Waterworth and Honey [10] reviewed 8 empirical studies of OHIS among older adults and discussed facilitators of and barriers to older adults' OHIS. However, the number of studies included in this review was limited, and it can hardly provide a comprehensive understanding of OHIS among older adults.

Gaps in the existing research indicate that a systematic scoping review on older adults' OHIS is necessary because it will not only enhance our knowledge of human information behaviors and practices but will also inform better health information system designs and ensure better information services for older adults. Motivated by the existing research gaps, this systematic scoping review examines the state of research on older adults' OHIS and reveals the types and sources of health information that the older adults seek, factors that influence older adults' OHIS, barriers to older adults' OHIS, and interventions that are available. The purpose of this systematic scoping review is to provide our readers with an overview of how OHIS among older adults has been studied and present implications for future research. It aims to answer the following questions:

1. What types of health information do older adults seek and where do they seek health information on the internet?

2. What are the factors that influence older adults' OHIS?

3. What are the barriers to older adults' OHIS?

4. How can we intervene and support older adults' OHIS?

Methods

Literature Search

This review follows the guidelines of the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) [11]. We were also inspired by the recommended framework for conducting systematic reviews in information-related fields by Okoli [12]. The bibliographic database search strategies were developed after consulting an academic librarian at the first author's university.

First, we searched the following databases: Web of Science; Cochrane Library database; PubMed; MEDLINE; CINAHL Plus; APA PsycINFO; Library and Information Science Source; Library, Information Science and Technology Abstracts; Psychology and Behavioral Sciences Collection; Communication & Mass Media Complete; ABI/INFORM; and ACM Digital Library. These databases were chosen because they cover the academic disciplines (eg, medicine, medical informatics, communication, psychology, and information and library science) that are most likely to study older adults' OHIS behaviors. Second, the search queries contained the following categories and keywords: people (older adults, elderly, aging, senior, seniors, older people, aged 60, aged 65), behavior (find, search, seek, access, retrieve), place (internet, online, web), object (information), and attribute (health, medicine, drug, nutrition, diet, wellness, illness). Specific queries were run in the topic, title, and abstract fields, depending on the database (see Multimedia Appendix 1). The initial search was performed in November 2020. Third, we captured additional articles using Google Scholar by tracking the citations and references in the articles found in the databases and in other relevant reviews. In addition, we supplemented relevant articles by searching Google Scholar directly. All the studies identified during the database searches were imported into the reference management software Zotero, and duplicates were removed.

Eligibility Criteria

We developed a series of inclusion and exclusion criteria to identify articles relating to older adults' OHIS behaviors. The inclusion criteria were as follows: (1) The articles should pertain to health-related contexts, including areas such as health, mental health, diet, and nutrition. (2) The article should describe OHIS behaviors (eg, general OHIS, selection and use of health information sources, and adoption and use of health information). (3) The article should focus on older adults (Note that although the search strategies indicated 2 commonly accepted lower age boundaries, 60 and 65 years, to identify older adults, it did not exclude other ways to describe the population); studies that clearly mentioned the population of older adults or contained explicit, equivalent claims were eligible. (4) The research should be empirically based. (5) The articles should have been published in a peer-reviewed journal or in conference proceedings. (6) When we identified more than 1 paper published by the same author on the same topic, we selected only the most recent one. (7) The articles should be written in English.

Our exclusion criteria were as follows: (1) The articles did not pertain to a health-related context. (2) The articles were not about OHIS behaviors; for instance, some articles focused only on general ICT use or adoption behaviors, were more concerned with technology-related rather than information-related issues or addressed only older adults' health literacy or eHealth literacy and did not investigate their OHIS. (3) The articles did not focus on older adults; we specifically excluded articles that treated age merely as a predictor or moderator in studying the OHIS of the general population, as it is evident that age influences people's OHIS behaviors. (4) The articles were not based on empirical research; this criterion helped eliminate opinion pieces, brief communications, editorial commentaries, and reviews. (5) The articles were not peer-reviewed (eg, a self-archived manuscript). (6) The articles were not written as full papers (eg, abstracts, posters, or letters). (7) The articles were not written in English.

Screening Procedure

The procedure for screening articles was based on the eligibility criteria. The initial search used database search strategies and identified 8047 publications. After duplicates were removed, the data set consisted of 5949 publications for screening.

The screening involved 3 stages. In the first stage, all the 3 authors reviewed the titles and abstracts of a sample of 300 articles from the search results, and then discussed and refined the screening criteria. In the second stage, we selected another 300 articles randomly from the search results as a test set. The feasibility criteria were verified independently by 2 of the authors (SS and MZ). Intercoder agreement (κ=0.816) indicated satisfactory reliability. Discrepancies were discussed and resolved by involving the third author (YZ), and the eligibility criteria were further refined accordingly. In the third stage, author MZ screened the remaining articles based on the eligibility criteria using the titles and abstracts, and author SS validated the results. Discrepancies were resolved by involving author YZ. The whole screening procedure resulted in 279 articles for full-text analysis.

To read and code the full-length articles downloaded from the databases, we used the MAXQDA 2020 software, which is designed for analyzing computer-assisted qualitative and mixed methods data, texts, and multimedia data. During the full-text analysis, we excluded 211 articles by applying the eligibility criteria. The remaining 68 articles were retained, and 8 more eligible articles were identified through citation tracking with the assistance of Google Scholar. In total, 75 articles were selected for the systematic scoping review.

Data Extraction and Analysis

We used Excel (Microsoft Corporation) to extract and record the basic information of the articles in the sample, including the author(s), title, publication year, publication name, and publication type (eg, journal vs conference). We used thematic content analysis in an iterative manner to identify the evidence regarding our research questions [13]. Several lists of codes were generated during 2 rounds of full-text coding procedures. In the first round, all the authors participated in the open and selective coding processes until a coding schema emerged and converged. In the second round, MZ coded the full texts by applying the coding schema, and SS validated all the codes. The intercoder reliability of the thematic content analysis reached 85%. Discrepancies were solved by involving YZ in the discussion.

Results

Basic Characteristics of the Included Articles

After screening, the final sample consisting of 75 articles was obtained, as shown in Figure 1. The articles were published between 1997 and 2020 (see Multimedia Appendix 2). Trend observations revealed that the number of publications in this subject area increased over time and that the OHIS of older adults began to receive considerable attention in the last 3 years (see Figure 2). The articles in the sample were mostly published after 2006 (n=69, 92%), which relates closely to the boom in social media. Of all the articles, 72 (96%) were published in journals, and the remaining 3 (4%) were published in conference proceedings. The articles originated from 17 countries (based on the first author's affiliations), with the top 3 being the United States (n=44, 58.67%), Australia (n=5, 6.67%), and China (n=4, 5.33%). The top 4 journals publishing these articles include the Journal of Medical Internet Research (n=8, 10.67%), Educational Gerontology (n=4, 5.33%), Journal of Health Communication (n=3, 4%), and Library & Information Science Research (n=3, 4%), indicating the multidisciplinary nature of the sample.

Figure 1. Screening procedure. ACM: Association for Computing Machinery; APA: American Psychological Association; CINAHL: Cumulative Index to Nursing and Allied Health Literature; ICT: information and communication technology.

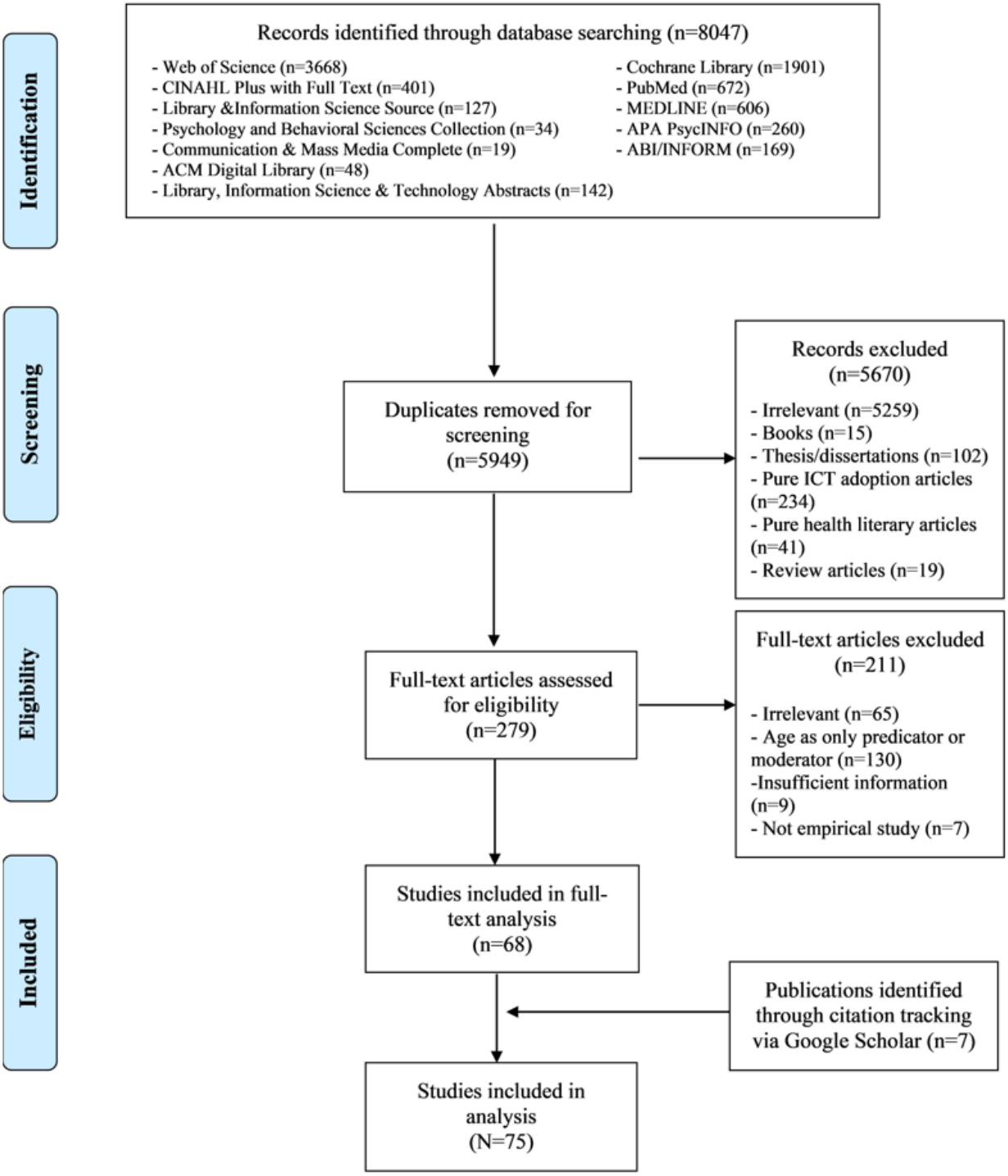
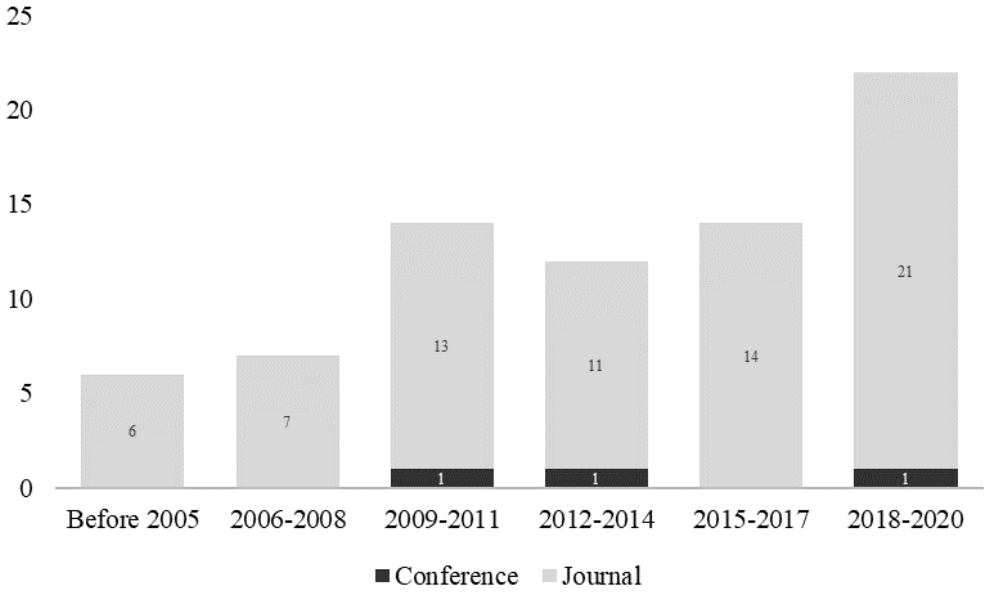


Figure 2. Distribution of publication years.



The systematic scoping review first investigated how the included 75 articles defined the target population of older adults. The cutoff ages for defining older adults were determined. More than half of the articles used samples of older adults aged above 60 years. Furthermore, 16 articles (21.33%) defined older adults as those aged 65 years and above, and 23 (30.67%) had cutoff ages ranging from 60 to 64 years. In addition, we noted some papers that defined the older adult group more loosely. For example, the cutoff age in 17 articles (22.67%) ranged from 50 to 54 years, and 14 articles (18.67%) used samples with minimum ages ranging from 55 to 59 years. Moreover, 5 of the articles (6.67%) did not specify precise age distributions.

The research methods varied across the 75 studies. Regarding methodological approaches, we found that 45 studies (60%) used quantitative approaches, 22 (29.33%) employed qualitative approaches, and 8 (10.67%) were based on mixed methods designs, using a combination of quantitative and qualitative methods. As for specific methods, surveys (n=28, 37.33%) and interviews or focus groups (n=25, 33.33%) were the primary methods used, followed by secondary data analysis (n=6, 8%) and experiments (n=4, 5.33%). In terms of data sources, most of the studies were based on primary data (n=65, 86.67%) and a few on secondary data (n=10, 13.33 %). Concerning the types of data, we found 59 studies (78.67%) based on cross-sectional data and 16 (21.33%) based on longitudinal data.

Internet-Based Health Information Types and Sources

Information types and information sources are 2 frequently reported aspects of information in OHIS studies [14]. For our analysis, we adapted the typologies of health information types from Kent et al [15] and Ramsey et al [16]. The results presented in Table 1 suggest that older adults often search the internet for information on specific diseases because they want to obtain a general idea of their condition before diagnosis or treatment so that they know what to expect and can be better prepared to face stressful situations [17]. The health problems mentioned in these 75 articles are mainly cancer (n=10, 13.33%), mental health problems (n=5, 6.67%), chronic conditions (n=4, 5.33%), and physical diseases (n=4, 5.33%). Aside from this disease information, the most frequently mentioned types of information are related to medication or treatment, nutrition or exercise, medical research, disease symptoms, and health promotion. Some articles mentioned that older adults also use the internet to seek support groups or interpersonal advice, health insurance information, health news, and health policy information. Of note is that more than half of the articles (n=40, 53.33%) used the umbrella term health information, without specifying any type of health information content. Furthermore, the types of content were not mutually exclusive. For example, a single article might mention more than 1 type of information (eg, older adults seeking information for cancer-related symptoms and medication).

Table 1. Types of health information mentioned in the articles (N=75).

Type of health information	Number of articles (n)
General health information	40
Specific diseases	23
Medication/treatment	21
Nutrition/exercise	13
Medical resource	12
Disease symptoms	9
Health promotion	8
Support groups/interpersonal advice	4
Health insurance	4
Health news/policies	3

Most of the articles in the sample (n=58, 77.33%) used the general internet to represent all the web-based sources of health information. Further, 26 articles (34.67%) described health websites as sources of internet-based health information for older adults; among these, the owners of the websites varied, consisting of educational, commercial, government, and nonprofit entities. Moreover, general search engines such as Google were the third most frequently mentioned sources in the studies (n=17, 22.67%), suggesting that older adults often use general search engines to start OHIS [18-20]. Further, 11 articles (14.67%) mentioned older adults' use of social media (eg, Facebook, Twitter) and blogs in OHIS. Only 3 articles (4%) addressed older adults' use of patient portals, and 2 articles (2.67%) were about older adults' use of mobile internet services. Table 2 shows the health information sources mentioned in the studies.

Table 2. Internet-based health information sources mentioned in the studies (N=75).

Source of internet-based health information	Number of articles (n)
General internet	58
Health websites (eg, WebMD, Mayo Clinic)	26
General search engines (eg, Google, Yahoo)	17
Social media/blogs (eg, Facebook, Twitter)	11
Patient portals	3
General mobile	2

Factors That Influence Older Adults' OHIS Behaviors

Among the 75 articles, 35 (46.67%) treated OHIS as a variable or construct. These articles quantitatively measured OHIS with various scales or proxy variables. Among them, 27 (36%) regarded OHIS as a dependent variable and explored the antecedents of older adults' OHIS. Further, 4 (5.33%) treated OHIS as an independent variable, and the remaining 4 (5.33%) treated OHIS as neither a dependent nor an independent variable but provided only descriptive analyses. Because the articles that employed quantitative approaches primarily concerned the antecedents of older adults' OHIS, we summarize the antecedents in Table 3.

We summarize the main influencing factors that appeared in the investigations. The antecedents of older adults' OHIS fall mainly into 2 categories, namely individual-related characteristics and source-related characteristics. Within the individual-related characteristics, 12 subcategories were observed, including demographics, anxiety, beliefs, attitudes, self-efficacy, personality, health status, medical history, health care service availability, source experience, health literacy and motivations. Among the source-related characteristics, credibility, usefulness, and ease of use were the 3 most frequently mentioned factors.

Table 3. Factors influencing older adults' online health information seeking behaviors.

Influencing Factors	Studies
Individual-related characteristics	
Demographics	
Socioeconomic status	[18,21-35]
Education	[18,21-41]
Gender	[18,21-27,29-34,36-43]
Marriage	[21,22,24,29,31,34,36,37,42,43]
Race/ethnicity	[23,29,30,33-35,37,38,40,44]
Place of residence	[24,30,36,41]
No. of children	[36]
Living with children	[33,35]
Anxiety	
ICT[a]-related anxiety	[18,42,45]
Disease-related fears	[31,46]
Perceived susceptibility	[21]
Beliefs	
External control	[45]
Internal locus of control	[28]
Fatalistic belief	[31,40]
Attitudes	
Attitudes on patient-doctor relationship	[18]
Reliance on and compliance with doctor's decisions	[18]
Attitudes on ICT use	[39,42]
Attitudes on internet-based health information	[38,45]
Attitudes on patient-doctor relationship	[18]
Self-efficacy	
Self-efficacy in health	[36]
Self-efficacy in learning	[24,42]
Self-efficacy in ICT use	[42,45]
Personality	
Big five	[36,42]
General values and life goals	[41]
Health status	
General health conditions	[18,21,28-30,32,38,41,42]
Physical health	[22,33-39]
Mental health	[31,34,36,37]
Chronic conditions	[22,30-36,43,44]
Medical history	
Personal medical history	[21]
Family medical history	[21,31]
Health care service availability	
Health care use	[34,36]
Health insurance status	[34]

Influencing Factors	Studies
Medical financial burden	[33]
Source experience	
Experience in internet use	[18,38,39,45]
Internet use frequency	[26,34,39]
Experience with online health information seeking	[29]
Experience in ICT use	[35,42]
Internet knowledge	[27]
Health literacy	
Health literacy	[24,33,43]
eHealth literacy	[27,29]
Motivations	
Health information needs	[18]
Health information orientation	[27]
Health information overload	[46]
Subjective norms	[39,45]
Source-related characteristics	
Credibility	
Trustworthiness	[28,38]
Relevance	[45]
Output quality	[45]
Result demonstrability	[45]
Usefulness	
Perceived usefulness of internet health information	[45]
Perceived usefulness of internet use	[28,39]
Perceived importance of health information	[28]
Ease of use	
Perceived ease of use of internet health information	[45]
Perceived ease of internet use	[39]
Computer playfulness	[45]
Perceived enjoyment	[45]

[a]ICT: information and communication technology.

Barriers to OHIS of Older Adults

Rather than treating OHIS as a variable, 40 of the 75 articles (53.33%) treated OHIS as a process. Of these studies, 29 (38.67%) explored the barriers that older adults encounter during OHIS. The results suggest that older adults may experience many barriers preventing successful OHIS, as shown in Table 4. In the prior studies, we identified 3 main types of barriers (ie, individual, social, and ICT), 11 subtypes, and 38 specific issues.

Table 4. Barriers to older adults' online health information seeking behavior.

Barrier types	Studies
Individual barriers	
Functional decline	
Vision impairment	[20,34]
Physical challenges (eg, back pain, knee injury)	[47,48]
Illness conditions	[32,35,36]
Low literacy	
English language literacy	[49,50]
Basic health knowledge	[51,52]
Digital literacy	[53,54]
Information literacy	[52,55,56]
Health literacy	[24,33,43]
eHealth literacy	[27,57]
Low self-efficacy	
Low efficacy and anxiety associated with computer use	[18,49,58,59]
Low efficacy in reading and learning	[49,60,61]
Low efficacy in OHIS[a]	[62,63]
Low efficacy in health information evaluation	[55,62]
Negative attitudes	
Attitude toward internet use	[39]
Attitude toward technology	[42]
Privacy concerns	[20,61,64]
Health beliefs	
External locus of control	[45]
Fatalistic beliefs	[31]
Social barriers	
Social stigmas	
Stigma of mental health problems	[65]
Stigma of sex-related health problems	[66]
Lack of social support	
Lack of informational support	[66,67]
Lack of organizational support (eg, health care services)	[17,50,68]
Lack of instrumental support (eg, instructions on computer use)	[57,65]
Lack of intergenerational support (eg, not living with children)	[49,69]
Lack of peer support (eg, hard to get support from friends)	[70,71]
ICT[b] barriers	
Lack of IT[c] infrastructure	
Lack of ICT devices	[29]
Low accessibility to medical records	[71]
Problematic information quality	
Misinformation	[64,72]
Conflicting health information	[73,74]

Barrier types	Studies
Irrelevant information	[65,73]
Information overload	
Overwhelming health information on the internet	[20,48,71]
Overwhelming extraneous information and pop-ups	[58,64,70]
Unsatisfactory user experiences	
Unsatisfactory interactivity and navigability	[75,76]
Unsuitable font sizes	[72,75]
Dense text and lack of visual elements	[76,77]
Confusing layouts	[51,72,75]
Insufficient ease of use	[39,45,78]
Frustrating user experiences	[51,56,59]

[a]OHIS: online health information seeking.

[b]ICT: information and communication technology.

[c]IT: information technology.

Regarding individual barriers, some studies found that older adults' OHIS could be hindered by age-related functional decline, including vision impairment, poor eye-hand coordination, physical challenges (eg, back pain), and illness. Moreover, some studies reported several aspects indicating low literacy among older adults that prevented effective OHIS, including limited English language skills, lack of basic health knowledge, limited digital literacy, undeveloped information literacy, and low health or eHealth literacy. Moreover, some studies found that older adults' perceptions of low self-efficacy regarding computer use, reading, learning, and evaluation of health information reduced their willingness toward OHIS. Other findings revealed that negative attitudes toward internet use or general technology and privacy concerns about using technology decreased older adults' intentions to search information on the internet. The results also revealed that beliefs regarding the external locus of the control of health care and fatalistic beliefs reduced older adults' active OHIS.

As for social barriers, studies suggested that older adults may have some social stigma concerning OHIS when it comes to mental and sex-related health problems. Moreover, older adults often report a lack of social support in their OHIS, including informational, organizational (eg, health care services), instrumental (eg, instructions on computer use), intergenerational (eg, support from children), and peer support (eg, support from friends).

In terms of ICT use, analysis of the studies revealed that many older adults do not possess information technology devices, and they reported low accessibility to medical records. Moreover, the quality of general health information on the internet is problematic. Older adults are likely to encounter misinformation, conflicting information, and irrelevant information during their OHIS. Furthermore, they often confront information overload when reading health information due to overwhelming amounts of irrelevant information or pop-ups. Moreover, older adults' OHIS may lead to some unpleasant and frustrating user experiences, such as unsatisfactory interactivity and navigability, unsuitable font sizes, dense text lacking visual elements, confusing layouts, and complicated site designs.

Interventions for Older Adults' OHIS

Given the abovementioned barriers, it is essential to provide older adults with additional support to facilitate their OHIS. We identified 11 studies (14.67%) among the 75 that used educational training programs to facilitate and intervene in older adults' OHIS, as shown in Table 5. Among these, 10 of the 11 studies provided offline workshops, and 1 conducted an online workshop. The offline workshops were conducted in community settings (eg, public libraries, schools, or medical centers) and included face-to-face instruction. We identified only 1 study that used an internet-based tutorial to improve older adults' ability to distinguish high-quality internet-based health forums from low-quality ones. Among the 11 articles, 9 described training programs with multiple sessions, each lasting 2 to 3 hours, and the duration of the programs varied from 1 to 4 months; the other 2 studies used 1-time training sessions.

Table 5. Interventions to support older adults' online health information seeking behaviors.

Study	Main objective	Intervention format	Intervention setting	Intervention evaluation measures
Malone et al [20]	To improve the health literacy skills of older adults	Educational program: Participants could attend every class offered at their library or could select the classes most appropriate to their personal needs and interests. No. of participants: 110	5 local libraries	Method: Pre- vs postsession surveys Qualitative analysis with descriptive statistics: Participants' confidence in their OHIS[a] increased, and the overall response to the program was positive.
Bertera et al [67]	To increase access to and use of 2 prominent health websites: MedlinePlus.gov and NIHSeniorHealth.gov	2-step training: (1) Training of internet navigators: 13 hours of basic training in computer skills over 13 weeks, plus a 4-hour specific training on 2 health websites and training on how to support peers during the process. No. of participants: 8 (2) Training of older adults living in affordable housing: 2-hour session on basic computer skills and use of 2 specific health websites. No. of participants: 42	A computer learning center located in the community	Method: Pre- vs posttest surveys, face-to-face interviews A significant improvement in the ability to use a computer or navigate the web was observed (P<.001). The average navigational skills self-efficacy score for health web sites (P<.001) and computers (P<.001) improved.
Chu et al [68]	To assist older adults with retrieving and evaluating health information resources on the internet	Educational program: 2-hour sessions once a week over 5 weeks. Partnering with Seniors for Better Health: Classes included 2 components, computer literacy and health information search strategies. No. of participants: 112	A computer lab offered at a facility of the YWCA[b] in Houston	Method: Pre- vs posttest surveys; survey conducted 6 weeks after training Participants experienced reduced computer anxiety and increased confidence and sense of self-efficacy when retrieving and evaluating internet-based health information (P<.001).
Campbell [79]	To improve the ability to locate health information	Workshops: 2-hour sessions once a week over 5 weeks The sessions used constructivist teaching techniques and self-directed learning. No. of participants: 70	A large suburban public library and 2 community centers for older adults	Method: Posttest interview Qualitative assessment by asking participants questions such as "Did your levels of participation in your health care change since you began using the internet?"
Campbell and Nolfi [80]	To teach older adults to access health care information on the internet	Workshops: 2-hour sessions once a week over 5 weeks No. of participants: 42 Follow-up survey 1 year after the workshops No. of participants: 27	A large suburban public library and 2 community centers for older adults	Method: Pre- vs. posttest surveys; survey 1 year after the training Statistically significant differences were found between baseline and 5-week follow-up results for MHLC[c] in males (P=.02) and females (P=.05), as well as for Krantz HOS[d] information seeking scores (P=.05).
Hoffman-Goetz et al [81]	To improve the internet search skills of adults aged 50 years and older	Workshops: 2-hour workshops once a month, over 4 months. The maximum number of participants per workshop was 15. Total No. of participants: 44	Public library with computer stations, led by a researcher, librarian, and university-based investigators	Method: Pre- vs posttest surveys Participants' search difficulty decreased after the workshops (P<.001). Participants' understanding of the internet improved after the workshops (P<.001).
Leung et al [82]	To improve basic skills for searching health information on the internet	Workshops: 3-hour training course The number of participants per workshop was 30. Total No. of participants: 88	Local university and company, instructed by nursing lecturer and students	Method: Postsession telephone interviews 1 month after the workshop Participants' confidence level in seeking health information was significantly associated with the level of satisfaction with the workshop (P<.001).
Campbell [83]	To improve health literacy skills among low-income, minority, and older adults	Workshops: 2-hour sessions once a week over 5 weeks No. of participants: 36	Computer labs in 2 low-income, minority residential buildings	Method: Pre- vs posttest surveys, survey 6 months after the training Participants experienced reduced anxiety concerning computers and increased confidence in locating health information.

Study	Main objective	Intervention format	Intervention setting	Intervention evaluation measures
Xie and Bugg [84]	To teach older adults to access and use high-quality internet-based health information	Educational program: 2-hour sessions twice a week over 4 weeks. The maximum number of participants per workshop was 7. Total No. of participants: 100	Public libraries	Method: Pre- vs posttest surveys Participants showed significantly reduced computer anxiety (P<.001), increased interest in computers (P=.001), and improved efficacy (P<.001) from pretraining to posttraining.
Chu and Mastel-Smith [85]	To enhance older adults' ability to grasp and manage health-related information retrieved from the internet and act accordingly	Educational program: 2-hour sessions once a week over 5 weeks. No. of participants: 12	A parish-sponsored, older adult leisure learning center	Method: Pre- vs posttest surveys; survey conducted 6 weeks after the training Participants experienced reduced anxiety, increased confidence, and a sense of self-efficacy at the end of the 5-week program and 6 weeks after program completion (P<.001).
Fink and Beck [86]	To improve the eHealth literacy of adults aged 50 years and older	Educational programs: 70 minutes to complete an educational online program and answer questions. No. of participants: 64	Internet-based setting	Method: Experimental group vs control group survey comparison Compared to the control group, the experimental group participants rated higher usability and learned more information on a new website.

[a]OHIS: online health information seeking.

[b]YWCA: Young Women's Christian Association.

[c]MHLC: multidimensional health locus of control.

[d]HOS: health opinion survey.

Further, 4 of the 11 programs were guided by established theories, models, or concepts (eg, the self-efficacy theory and the health belief model). All the studies involved some form of evaluation, including postsession surveys or interviews, pre- versus postintervention comparisons, and experimental versus control group comparisons. In addition, 5 studies evaluated the effectiveness of the intervention outcomes from a longitudinal perspective over a period ranging from 1 month to 1 year to the competence of the program. Among all the studies, 9 statistically assessed the effects of the intervention. Measures varied across the studies; these included opinions from surveys on the internet, self-efficacy in seeking health information, and anxiety regarding computer use. All the articles reported some positive outcomes of the intervention programs.

Discussion

Principal Findings

This systematic scoping review provides an overview of OHIS behaviors among older adults, as shown in Figure 3. Overall, the findings of this paper reveal core elements of OHIS among older adults. First, the types and sources of health information that older adults search for were clearly presented. Then, a portion of the studies explored the main factors influencing older adults' OHIS behaviors, which can be categorized as individual-related and source-related characteristics. Then, we identified the barriers to OHIS behavior in older adults from existing literature, including individual barriers, social barriers, and ICT barriers. Finally, this paper provides an in-depth analysis of the interventions mentioned in some of the included papers to support OHIS behaviors among older adults. We believe that the framework of this paper can, to some extent, help researchers to better position their research objectives in future studies so that the objectives correspond to specific dimensions for in-depth empirical investigation.

Regarding the first research question, the results show that older adults sought various types of health information on the internet, including information about specific diseases, medication and treatment, nutrition and exercise, medical resources, disease symptoms, health promotion, support groups and interpersonal advice, health insurance information, and health news or policies. The information sources included health websites, general search engines, social media and blogs, patient portals, and mobile devices. The types of health information sought differed from those that interest young people. According to a recent systematic review [87], adolescents and youths (<24 years) search the internet for daily health-related issues, physical and psychological well-being, sexual health, social problems, and culturally sensitive topics. Compared to the adolescent and youth population, older adults tend to search more for disease-related health information topics.

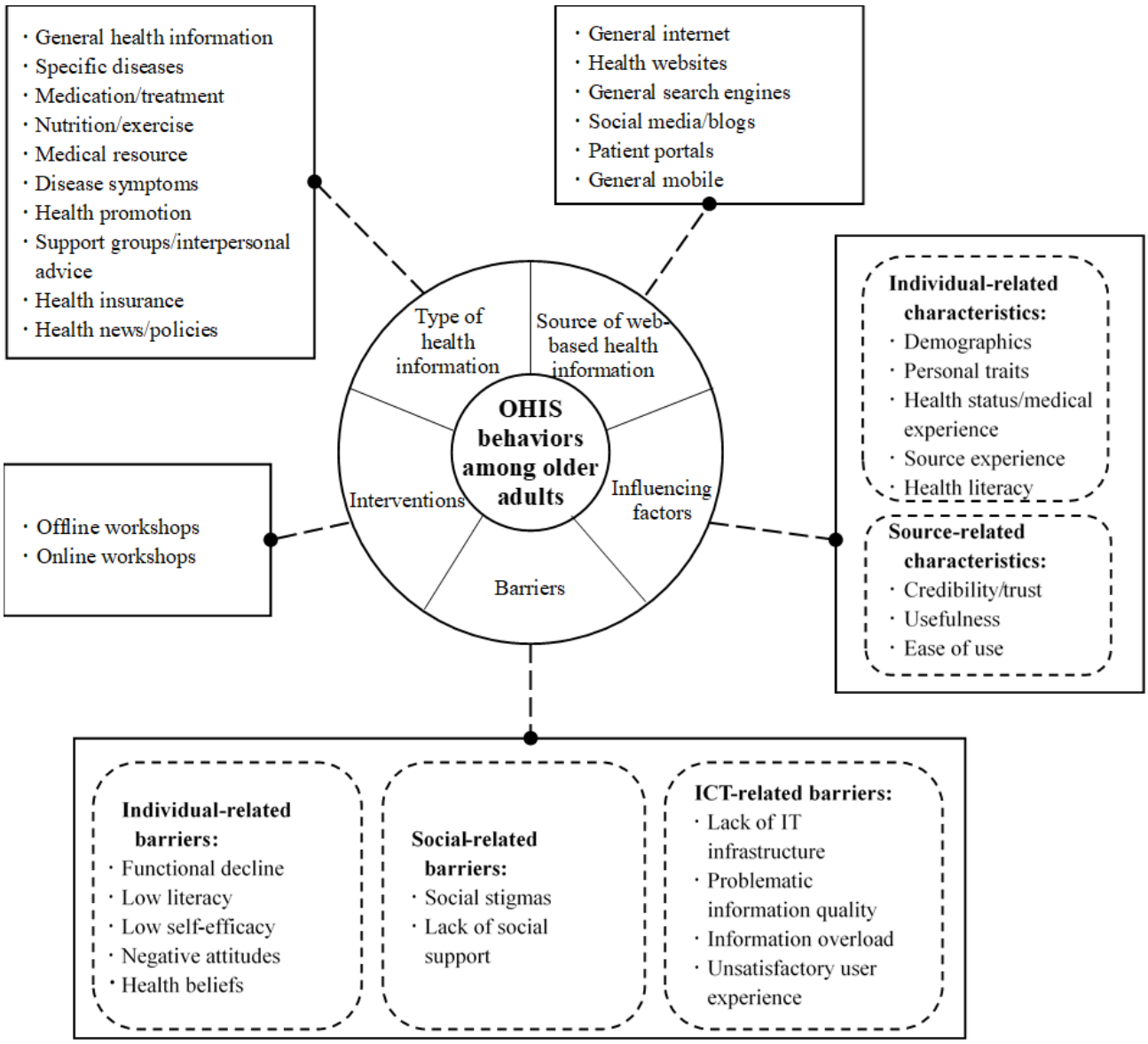
As for the second research question, the results point to 2 main types of factors influencing older adults' OHIS: individual-related characteristics and source-related characteristics. The individual-related characteristics include demographics, anxiety, beliefs, attitudes, self-efficacy, personality, health status, medical history, health care service availability, source experience, health literacy, and motivations. Among the source-related characteristics, credibility, usefulness, and trust were the 3 factors most frequently mentioned in the studies. We noted that the primary factors influencing older adults' OHIS differ from those influencing young adults. A systematic review of studies investigating young adults' (<24 years) OHIS [87] revealed that the most frequently mentioned influencing factors were gender, age, educational status, emotional characteristics, engagement in risky behaviors, and eHealth literacy.

The results for the third research question reveal that older adults might encounter 3 types of barriers during their OHIS, including

individual barriers (eg, low literacy), social barriers (eg, social stigmas), and ICT-related barriers (eg, lack of ICT devices). These barriers may hinder effective OHIS behaviors of older adults. The results suggest some differences from the findings on young adults' OHIS. For the adolescent and youth population (<24 years), the main barriers to OHIS include online privacy and concerns about information credibility [87]. Although some studies report low health literacy among adolescents [88], older adults seem to have more difficulties in this respect than adolescents [89,90].

As for the fourth research question, the review found that many intervention programs have been created to support older adults' OHIS; they primarily use educational training workshops in offline and online formats. Most training programs contained multiple sessions, with each session lasting 2 to 3 hours; the duration of the programs varied from 1 to 4 months, and all the programs reported at least some positive effects in support of older adults' OHIS.

Figure 3. Overview of principal findings. ICT: information and communication technology; IT: information technology; OHIS: online health information seeking.



Implications for Future Research

Overall, this systematic scoping review identified the need for more in-depth research on older adults' OHIS. As can be seen from the aforementioned evidence, a subset of studies have treated OHIS as a variable or construct and focused on exploring the factors influencing OHIS in older adults. Other studies treat OHIS as a process and investigate how the older adults search the internet for health information. However, given the complexity of the health conditions of older people and a projected future intensification of information overload, older adults will encounter more serious problems when searching for health information on the internet, such as how to select from among multimodal information sources, how to express health information needs, and how to evaluate health misinformation. Considering the growing population of older adults, the importance of internet-based information seeking for overall public health, and the lack of best practices, more research on this topic is needed. In this section, we propose

several directions for future research based on gaps identified in the review.

Investigations on the OHIS Behavior of Older Adults Above 85 Years

With the accelerating pace of global aging, the population of older adults is steadily growing. Instead of classifying the large population of older adults as one group, researchers are advocating for a more precise segmentation of this population, such as the youngest-old (65 to 74 years), middle-old (75 to 84 years), and oldest-old groups (above 85 years) [91]. Regarding OHIS, the age distribution of the samples in this systematic scoping review indicates that the exploration of OHIS by the oldest-old group is very limited [92]. Most articles included in this review have focused on the youngest- and middle-old groups [30], whereas there is a lack of research on the health information needs and behaviors of the oldest-old group. Future OHIS research can be appropriately skewed toward the oldest-old group to consider the physiological and psychological characteristics, the unique information needs, and explore the influences, processes, and health outcomes of the OHIS of this group more empirically within the framework of everyday information mastering [93].

Conducting More Longitudinal, Action, and Mixed Methods Research

As for research methods, most current studies use cross-sectional data collection methods and pay little attention to longitudinal approaches. In future, more consideration can be given to the adoption of longitudinal methods, such as the experience sampling method and the ethnographic approach. In particular, for intervention studies on OHIS behaviors in older adults, educational training programs with long time spans could provide data to improve OHIS performance and the health literacy of older adults. More participatory action research at the community level would enrich the network of actors in OHIS for older adults and engage more participants, thereby promoting interdisciplinary and collaborative health information practices in this population. In addition, future studies might consider more mixed methods approaches to leverage the advantages of qualitative and quantitative approaches and triangulate primary data with secondary data. Existing mixed methods studies have been conducted mainly based on quantitative questionnaire analyses as well as qualitative focus groups, and a richer mix of methods is to be further explored for this topic in future. Finally, as prior studies have relied heavily on self-reported data, future studies could consider more behavioral data using methods such as eye-tracking and electroencephalograms.

Elaboration on Mobile Context and Cross-platform Scenario of Older Adults' OHIS

Information types and information sources are the essential contextual factors in OHIS [94-96]. However, this review found that most studies on older adults' OHIS do not clearly explain what health-related information was involved or from where the information was gathered. In terms of information types, current studies mainly focus on searches for disease and treatment information. More studies are needed to address other types of health information that older adults might seek, such as information on environmental health and disease prevention.

Regarding information sources, studies are needed to investigate older adults' use of mobile devices for OHIS. With the development of the mobile internet and the internet of things, OHIS scenarios for older adults are changing. Mobile device–based health information access can more effectively meet the health information needs of older adults, facilitate daily health monitoring and self-tracking, and improve context-driven, health-related decision-making among older adults. For example, increasing numbers of older adults are seeking health information on their smart phones through short video apps like TikTok [97,98]. Furthermore, in addition to searching for health information on their mobile devices, increasing numbers of older adults are using mobile social apps to create content [99]. Future research could focus more on the relationship between OHIS and health-related content generation by older adults.

In addition, further exploration of complicated OHIS scenarios is needed. For example, with the popularity of wearable devices and the development of various health-related vertical search platforms, a portion of the older adult population with higher information literacy will become more proficient at searching for a full range of health information using various smart devices and immersive technologies [100], such as interacting with information through voice recognition and gesture control. Thus, explorations of cross-platform and cross-device seeking behaviors in OHIS by older adults are needed. Meanwhile, in addition to active information seeking, more types of seeking behaviors, such as passive exposure, information encountering, and surrogate health information seeking [101,102], deserve attention and further investigation. In particular, the influences and positive outcomes of searching as learning during OHIS by older adults is a topic worth exploring.

Facilitating Older Adults' OHIS by Explicating Technology Affordance

This review revealed that current research on factors influencing OHIS in older adults focuses more on demographic issues and individual-related characteristics than on source-related factors. In recent years, increased emphasis has been placed on aging-friendly designs in human-computer interaction [103], and the user experience–oriented design of various social apps and smart devices is centered on the needs and behavioral preferences of older adults, with an interest in meeting their personalized requirements. We believe that the affordance of technology in aging-friendly design is also a highly influential factor for promoting OHIS in older adults. It would be fruitful to integrate the uses and gratifications theory with the affordance lens to better promote the positive impact of new media platforms on older adults' information-seeking behaviors [104,105]. More attention needs to be placed on the ease of use, usability, and sociability of aging-friendly information sources and information systems. In particular, in the upcoming human-centered artificial intelligence era, older people's perception of the trustworthiness of multimodal information sources and their trust in algorithm-based content recommendations will continue to change. Therefore, the age-appropriate design of OHIS needs to constantly break away

from stereotypes of older people and re-establish a more adaptive mental model. The lens of the affordance theory could be applied to help situate OHIS for older adults in the context of information practices, promoting deep reflection on the interaction of actors with sociocultural environments and on the mediated nature of technology [106]. For instance, an OHIS platform should provide rich technology affordances for older adults and provide targeted support for active health information access, information encounters, and information avoidance problems in different sociocultural environments. Future research could focus more on how technology affordance can better mediate older adults' OHIS gratification by attempting to build a more detailed affordance typology [107]—such as handling, effecter, and motivational affordances—to measure older adults' gratifications for OHIS using social media.

Promoting and Measuring the Performance of OHIS Interventions for Older Adults

The results show that older adults encounter many barriers in OHIS; thus, many intervention programs have been created to support their searching. However, current intervention programs still leave considerable room for improvement. First, current educational training programs are generally small-scale ones, making it difficult to reach a wide group of older adults; most programs are offline workshops, and there are few internet-based programs. Future OHIS interventions for older adults need to offer more technology-mediated web-based programs and provide richer formats than workshops and tutorials, such as distance education for older adults using gamification and immersive technology. Moreover, most current intervention programs operate in the United States; older adults living in less developed countries or areas received less attention. Future studies on OHIS in older adults must involve more trans- and cross-national, or regional and cross-cultural comparative studies to further explore the influence of sociocultural factors on older adults' OHIS behaviors. We also recommend that more information and communication technology for development (known as “ICT4D”) projects focus on upgrading OHIS and improving the same for older adults [108], thereby better promoting health literacy and health mobility for older adults in developing countries and regions.

In particular, researchers need to draw more on the design science research paradigm. Design science research is an innovative and often iterative problem-solving process that builds and evaluates artifacts [109]. In our research context, the purposeful artifacts could be search systems, training courses, workshops, tutorials, or citizen science programs. In the building phase of artifact development, most units of analysis relate to offline workshops and neglect other types of artifacts. It is also noteworthy that current intervention studies lack a theoretical lens, and only a few studies have designed interventions based on theoretical foundations. Future interventions for older adults' OHIS need to embrace the theoretical considerations that design science research has been advocating [110]. In the evaluation phase of artifact development, current studies lack long-term assessments of intervention effects. Future studies should consider more participatory action research to iteratively test the effects of OHIS interventions on older adults and select some specific health domains—such as chronic diseases, cancer, and mental health—for attempting to verify the actual effects of OHIS interventions on information literacy, health literacy, and health outcomes of older adults. In addition, future studies could contemplate providing various forms of support based on the perspectives of older users, allowing them to participate in the project design process and thus help them overcome search barriers.

Limitations

This systematic scoping review has several limitations. The first one is in terms of search sources. Owing to the interdisciplinary nature of OHIS research in older adults, although we tried to search multiple databases using relevant keywords and consulted academic librarians to improve our search strategy, it was nevertheless inevitable that some literature would be missed, especially relevant research in unofficially published conference proceedings. The backward and forward strategy can be further used to expand the literature search sources in future [111]. Second, in terms of the literature type, this review mainly focuses on empirical studies, whereas some opinion papers, descriptive cases, and short communications on OHIS for older adults were excluded from our literature pool, and some complementary analyses of such nonresearch articles can be conducted in future. Finally, in terms of the analytical approach for searching literature, this study did not conduct a comparative chronological analysis of the literature in different periods, which to a certain extent could not fully reveal the impact of technological and sociocultural changes on older adults’ OHIS behavior. In future, the introduction of knowledge graphs can be considered to map the themes of the literature at different stages.

Conclusions

This review provides an overview of how older adults’ OHIS has been studied. It reveals that older adults search for various types of health information on the internet using different types of web-based sources and that their OHIS is jointly influenced by source-related and individual-related factors. Their difficulties in searching arise from individual, social, and ICT-related barriers. Some educational intervention programs that support older adults’ OHIS have been initiated in the form of web-based and offline workshops. Furthermore, the review reveals that the topic of older adults’ OHIS is understudied, although the number of studies is increasing. Nevertheless, more studies are needed to understand the problems associated with older adults’ interactions with health information and better support them in their decision-making when they are searching for medical and health information on the internet. Based on the findings of the review, the authors propose several objectives for future research.

Acknowledgments

We thank the anonymous reviewers and editors for their helpful comments on the manuscript. The study was supported by the National Science Foundation of China (grants 72074112, 72174083, and 71774083).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Database search strategies.
[DOCX File , 14 KB - jmir_v24i2e34790_app1.docx]

Multimedia Appendix 2
Overview of included studies.
[DOCX File , 61 KB - jmir_v24i2e34790_app2.docx]

References

1. Chi Y, He D, Jeng W. Laypeople's source selection in online health information-seeking process. J Assoc Inf Sci Technol 2020 Feb;71(12):1484-1499. [doi: 10.1002/asi.24343]
2. Pian W, Song S, Zhang Y. Consumer health information needs: a systematic review of measures. Inf Process Manag 2020 Mar;57(2):102077 [FREE Full text] [doi: 10.1016/j.ipm.2019.102077]
3. Cline R, Haynes K. Consumer health information seeking on the internet: the state of the art. Health Educ Res 2001 Dec;16(6):671-692. [doi: 10.1093/her/16.6.671] [Medline: 11780707]
4. Kovner CT, Mezey M, Harrington C. Who cares for older adults? Workforce implications of an aging society. Health Aff (Millwood) 2002 Sep;21(5):78-89. [doi: 10.1377/hlthaff.21.5.78] [Medline: 12224911]
5. Medlock S, Eslami S, Askari M, Arts DL, Sent D, de Rooij SE, et al. Health information-seeking behavior of seniors who use the internet: a survey. J Med Internet Res 2015 Jan;17(1):e10 [FREE Full text] [doi: 10.2196/jmir.3749] [Medline: 25574815]
6. Czaja SJ. Can technology empower older adults to manage their health? Generations 2015 Mar;39(1):46-51 [FREE Full text]
7. Yoon H, Jang Y, Vaughan PW, Garcia M. Older adults' internet use for health information: digital divide by race/ethnicity and socioeconomic status. J Appl Gerontol 2020 Jan;39(1):105-110. [doi: 10.1177/0733464818770772] [Medline: 29661052]
8. Hunsaker A, Hargittai E. A review of internet use among older adults. New Media Soc 2018 Jul;20(10):3937-3954. [doi: 10.1177/1461444818787348]
9. Chang C, Huang M. Antecedents predicting health information seeking: a systematic review and meta-analysis. Int J Inf Manag 2020 Oct;54:102115. [doi: 10.1016/j.ijinfomgt.2020.102115]
10. Waterworth S, Honey M. On-line health seeking activity of older adults: an integrative review of the literature. Geriatr Nurs 2018 May;39(3):310-317. [doi: 10.1016/j.gerinurse.2017.10.016] [Medline: 29198622]
11. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, Tunçalp, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct;169(7):467-473. [doi: 10.7326/m18-0850]
12. Okoli C. A guide to conducting a standalone systematic literature review. CAIS 2015;37(43):879-910. [doi: 10.17705/1CAIS.03743]
13. Wildemuth BM, editor. Applications of Social Research Methods to Questions in Information and Library Science. Santa Barbara, CA: ABC-CLIO; 2016.
14. Anker AE, Reinhart AM, Feeley TH. Health information seeking: a review of measures and methods. Patient Educ Couns 2011 Mar;82(3):346-354. [doi: 10.1016/j.pec.2010.12.008] [Medline: 21239134]
15. Kent EE, Arora NK, Rowland JH, Bellizzi KM, Forsythe LP, Hamilton AS, et al. Health information needs and health-related quality of life in a diverse population of long-term cancer survivors. Patient Educ Couns 2012 Nov;89(2):345-352 [FREE Full text] [doi: 10.1016/j.pec.2012.08.014] [Medline: 23021856]
16. Ramsey I, Corsini N, Peters MD, Eckert M. A rapid review of consumer health information needs and preferences. Patient Educ Couns 2017 Sep;100(9):1634-1642. [doi: 10.1016/j.pec.2017.04.005] [Medline: 28442155]
17. Xie B. Older adults' health information wants in the internet age: implications for patient–provider relationships. J Health Commun 2009 Sep;14(6):510-524. [doi: 10.1080/10810730903089614] [Medline: 19731124]
18. McMillan SJ, Macias W. Strengthening the safety net for online seniors: factors influencing differences in health information seeking among older internet users. J Health Commun 2008 Dec;13(8):778-792. [doi: 10.1080/10810730802487448] [Medline: 19051113]

19. Huisman M, Joye S, Biltereyst D. Searching for health: Doctor Google and the shifting dynamics of the middle-aged and older adult patient-physician relationship and interaction. J Aging Health 2020 Oct;32(9):998-1007. [doi: 10.1177/0898264319873809] [Medline: 31517558]
20. Malone T, Jo P, Clifton S. Perceived eHealth literacy and information behavior of older adults enrolled in a health information outreach program. J Consum Health Internet 2017 Jun;21(2):137-147. [doi: 10.1080/15398285.2017.1300040]
21. Meischke H, Eisenberg M, Rowe S, Cagle A. Do older adults use the internet for information on heart attacks? Results from a survey of seniors in King County, Washington. Heart Lung 2005 Jan;34(1):3-12. [doi: 10.1016/j.hrtlng.2004.06.006] [Medline: 15647729]
22. Tak SH, Hong SH. Use of the internet for health information by older adults with arthritis. Orthop Nurs 2005 Mar;24(2):134-138. [doi: 10.1097/00006416-200503000-00010] [Medline: 15902012]
23. Cresci MK, Jarosz PA, Templin TN. Are health answers online for older adults? Educ Gerontol 2012 Jan;38(1):10-19. [doi: 10.1080/03601277.2010.515890]
24. Lam MK, Lam LT. Health information-seeking behaviour on the Internet and health literacy among older Australians. Electron J Health Inform 2012 Oct:e15 [FREE Full text]
25. Zajac IT, Flight IHK, Wilson C, Turnbull D, Cole S, Young G. Internet usage and openness to internet-delivered health information among Australian adults aged over 50 years. Australas Med J 2012 Jun;5(5):262-267 [FREE Full text] [doi: 10.4066/AMJ.2012.1065] [Medline: 22848321]
26. Chaudhuri S, Le T, White C, Thompson H, Demiris G. Examining health information-seeking behaviors of older adults. Comput Inform Nurs 2013 Nov;31(11):547-553 [FREE Full text] [doi: 10.1097/01.NCN.0000432131.92020.42] [Medline: 23974574]
27. Sheng X, Simpson PM. Seniors, health information, and the internet: motivation, ability, and internet knowledge. Cyberpsychol Behav Soc Netw 2013 Oct;16(10):740-746. [doi: 10.1089/cyber.2012.0642] [Medline: 23679569]
28. Sheng X, Simpson PM. Health care information seeking and seniors: determinants of internet use. Health Mark Q 2015 Mar;32(1):96-112. [doi: 10.1080/07359683.2015.1000758] [Medline: 25751321]
29. Tennant B, Stellefson M, Dodd V, Chaney B, Chaney D, Paige S, et al. eHealth literacy and Web 2.0 health information seeking behaviors among baby boomers and older adults. J Med Internet Res 2015 Mar;17(3):e70 [FREE Full text] [doi: 10.2196/jmir.3992] [Medline: 25783036]
30. Hong YA, Cho J. Has the digital health divide widened? Trends of health-related internet use among older adults from 2003 to 2011. J Gerontol B Psychol Sci Soc Sci 2017 Sep;72(5):856-863. [doi: 10.1093/geronb/gbw100] [Medline: 27558403]
31. Leung DY, Chow TT, Wong EM. Cancer-related information seeking and scanning behaviors among older Chinese adults: examining the roles of fatalistic beliefs and fear. Geriatrics (Basel) 2017 Dec;2(4):38 [FREE Full text] [doi: 10.3390/geriatrics2040038] [Medline: 31011048]
32. Burns P, Jones SC, Caputi P, Iverson D. Are older Australians with chronic diseases online? Health Promot J Austr 2018 Apr;29(1):72-78. [doi: 10.1002/hpja.5] [Medline: 29700931]
33. Oh YS, Choi EY, Kim YS. Predictors of smartphone uses for health information seeking in the Korean elderly. Soc Work Public Health 2017 Dec;33(1):43-54. [doi: 10.1080/19371918.2017.1391150] [Medline: 29257932]
34. Choi NG, DiNitto DM, Lee OE, Choi BY. Internet and health information technology use and psychological distress among older adults with self-reported vision impairment: case-control study. J Med Internet Res 2020 Jun;22(6):e17294 [FREE Full text] [doi: 10.2196/17294] [Medline: 32490851]
35. Sedrak MS, Soto-Perez-De-Celis E, Nelson RA, Liu J, Waring ME, Lane DS, et al. Online health information-seeking among older women with chronic illness: analysis of the women's health initiative. J Med Internet Res 2020 Apr;22(4):e15906 [FREE Full text] [doi: 10.2196/15906] [Medline: 32271152]
36. Flynn KE, Smith MA, Freese J. When do older adults turn to the internet for health information? Findings from the Wisconsin Longitudinal Study. J Gen Intern Med 2006 Sep;21(12):1295-1301 [FREE Full text] [doi: 10.1111/j.1525-1497.2006.00622.x] [Medline: 16995892]
37. Nguyen GT, Shungu NP, Niederdeppe J, Barg FK, Holmes JH, Armstrong K, et al. Cancer-related information seeking and scanning behavior of older Vietnamese immigrants. J Health Commun 2010 Nov;15(7):754-768 [FREE Full text] [doi: 10.1080/10810730.2010.514034] [Medline: 21104504]
38. Zulman DM, Kirch M, Zheng K, An LC. Trust in the internet as a health resource among older adults: analysis of data from a nationally representative survey. J Med Internet Res 2011 Feb;13(1):e19 [FREE Full text] [doi: 10.2196/jmir.1552] [Medline: 21324832]
39. Wong CKM, Yeung DY, Ho HCY, Tse K, Lam C. Chinese older adults' internet use for health information. J Appl Gerontol 2014 Apr;33(3):316-335. [doi: 10.1177/0733464812463430] [Medline: 24717738]
40. Lumpkins CY, Mabachi N, Lee J, Pacheco C, Greiner KA, Geana M. A prescription for internet access: appealing to middle-aged and older racial and ethnic minorities through social network sites to combat colorectal cancer. Health Commun 2017 Jul;32(7):916-920. [doi: 10.1080/10410236.2016.1195679] [Medline: 27435103]
41. Weber W, Reinhardt A, Rossmann C. Lifestyle segmentation to explain the online health information-seeking behavior of older adults: representative telephone survey. J Med Internet Res 2020 Jun;22(6):e15099 [FREE Full text] [doi: 10.2196/15099] [Medline: 32530433]

42. Parida V, Mostaghel R, Oghazi P. Factors for elderly use of social media for health-related activities. Psychol Mark 2016 Nov;33(12):1134-1141. [doi: 10.1002/mar.20949]
43. Kim SH, Utz S. Association of health literacy with health information-seeking preference in older people: a correlational, descriptive study. Nurs Health Sci 2018 Sep;20(3):355-360. [doi: 10.1111/nhs.12413] [Medline: 29488295]
44. Li Y, Polk J, Plankey M. Online health-searching behavior among HIV-seropositive and HIV-seronegative men who have sex with men in the Baltimore and Washington, DC area. J Med Internet Res 2013 May;15(5):e78 [FREE Full text] [doi: 10.2196/jmir.2479] [Medline: 23644412]
45. Chang SJ, Im EO. A path analysis of internet health information seeking behaviors among older adults. Geriatr Nurs 2014 Mar;35(2):137-141. [doi: 10.1016/j.gerinurse.2013.11.005] [Medline: 24332965]
46. Jensen JD, Liu M, Carcioppolo N, John KK, Krakow M, Sun Y. Health information seeking and scanning among US adults aged 50-75 years: testing a key postulate of the information overload model. Health Informatics J 2017 Jun;23(2):96-108 [FREE Full text] [doi: 10.1177/1460458215627290] [Medline: 26905079]
47. Harrod M. "I Have to Keep Going": Why some older adults are using the internet for health information. Ageing Int 2011 Jun;36(2):283-294. [doi: 10.1007/s12126-010-9090-z]
48. Chung J, Gassert CA, Kim HS. Online health information use by participants in selected senior centres in Korea: current status of internet access and health information use by Korean older adults. Int J Older People Nurs 2011 Dec;6(4):261-271. [doi: 10.1111/j.1748-3743.2010.00238.x] [Medline: 21631875]
49. Goodall KT, Newman LA, Ward PR. Improving access to health information for older migrants by using grounded theory and social network analysis to understand their information behaviour and digital technology use. Eur J Cancer Care (Engl) 2014 Nov;23(6):728-738. [doi: 10.1111/ecc.12241] [Medline: 25250535]
50. Agyemang-Duah W, Arthur-Holmes F, Peprah C, Adei D, Peprah P. Dynamics of health information-seeking behaviour among older adults with very low incomes in Ghana: a qualitative study. BMC Public Health 2020 Jun;20(1):928 [FREE Full text] [doi: 10.1186/s12889-020-08982-1] [Medline: 32539693]
51. Huang M, Hansen D, Xie B. Older adults' online health information seeking behavior. In: iConference '12: Proceedings of the 2012 iConference. New York, United States: Association for Computing Machinery; 2012 Presented at: 2012 iConference; 7-10 February, 2012; Toronto, Canada p. 338-345. [doi: 10.1145/2132176.2132220]
52. Sanchiz M, Chin J, Chevalier A, Fu W, Amadieu F, He J. Searching for information on the web: impact of cognitive aging, prior domain knowledge and complexity of the search problems. Inf Process Manage 2017 Jan;53(1):281-294 [FREE Full text] [doi: 10.1016/j.ipm.2016.09.003]
53. Czaja SJ, Sharit J, Hernandez MA, Nair S, Loewenstein D. Variability among older adults in Internet health information-seeking performance. Gerontechnology 2010;9(1):46-55. [doi: 10.4017/gt.2010.09.01.004.00]
54. Pálsdóttir Á. Analysing digital information seeking and quality evaluation of sources about healthy lifestyle among senior citizens in Iceland. RRBSI 2018;14(2):35-45. [doi: 10.26660/rrbsi.2018.14.2.35]
55. Robertson-Lang L, Major S, Hemming H. An exploration of search patterns and credibility issues among older adults seeking online health information. Can J Aging 2011 Dec;30(4):631-645. [doi: 10.1017/S071498081100050X] [Medline: 22085455]
56. Wu D, Li Y. Online health information seeking behaviors among Chinese elderly. Libr Inf Sci Res 2016 Jul;38(3):272-279. [doi: 10.1016/j.lisr.2016.08.011]
57. Marcu A, Muller C, Ream E, Whitaker KL. Online information-seeking about potential breast cancer symptoms: capturing online behavior with an internet browsing tracking tool. J Med Internet Res 2019 Feb;21(2):e12400 [FREE Full text] [doi: 10.2196/12400] [Medline: 30724741]
58. Taha J, Sharit J, Czaja S. Use of and satisfaction with sources of health information among older internet users and nonusers. Gerontologist 2009 Oct;49(5):663-673 [FREE Full text] [doi: 10.1093/geront/gnp058] [Medline: 19741112]
59. Zapata AML, Beaudreau SA, O'Hara R, Bereknyei Merrell S, Bruce J, Garrison-Diehn C, et al. Information-seeking about anxiety and perceptions about technology to teach coping skills in older veterans. Clin Gerontol 2018 Oct;41(4):346-356. [doi: 10.1080/07317115.2017.1359716] [Medline: 28967837]
60. Gollop CJ. Health information-seeking behavior and older African American women. Bull Med Libr Assoc 1997 Apr;85(2):141-146 [FREE Full text] [Medline: 9160150]
61. Huvila I, Enwald H, Eriksson-Backa K, Hirvonen N, Nguyen H, Scandurra I. Anticipating ageing: older adults reading their medical records. Inf Process & Manag 2018 May;54(3):394-407. [doi: 10.1016/j.ipm.2018.01.007]
62. Mayoh J, Todres L, Bond CS. Exploring the online health information seeking experiences of older adults. Indo-Pac J Phenomenol 2011 Oct;11(2):1-13. [doi: 10.2989/ipjp.2011.11.2.3.1162]
63. Ortiz-Dowling EM, Der Ananian C, Larkey LK, Hooker SP. Health-seeking behaviors and health information gathering in older Mexican American males. Psychol Men Masc 2019 Oct;20(4):564-574. [doi: 10.1037/men0000194]
64. Choi W. Older adults' health information behavior in everyday life settings. Libr Inf Sci Res 2019 Oct;41(4):100983. [doi: 10.1016/j.lisr.2019.100983]
65. Moult A, Burroughs H, Kingstone T, Chew-Graham CA. How older adults self-manage distress - does the internet have a role? A qualitative study. BMC Fam Pract 2018 Nov;19(1):185 [FREE Full text] [doi: 10.1186/s12875-018-0874-7] [Medline: 30497414]

66. Lyons A, Mikolajczak G, Heywood W, Fileborn B, Minichiello V, Hinchliff S, et al. Sources of information-seeking on sexually transmitted infections and safer sex by older heterosexual Australian men and women. Educational Gerontology 2018 Feb;44(2-3):186-195. [doi: 10.1080/03601277.2018.1433989]
67. Bertera EM, Bertera RL, Morgan R, Wuertz E, Attey AMO. Training older adults to access health information. Educational Gerontology 2007 May;33(6):483-500. [doi: 10.1080/03601270701328250]
68. Chu A, Huber J, Mastel-Smith B, Cesario S. "Partnering with Seniors for Better Health": computer use and Internet health information retrieval among older adults in a low socioeconomic community. J Med Libr Assoc 2009 Jan;97(1):12-20 [FREE Full text] [doi: 10.3163/1536-5050.97.1.003] [Medline: 19159002]
69. Altizer KP, Grzywacz JG, Quandt SA, Bell R, Arcury TA. A qualitative analysis of how elders seek and disseminate health information. Gerontol Geriatr Educ 2014;35(4):337-353 [FREE Full text] [doi: 10.1080/02701960.2013.844693] [Medline: 24188253]
70. Chang L, Basnyat I, Teo D. Seeking and processing information for health decisions among elderly Chinese Singaporean women. J Women Aging 2014 Jun;26(3):257-279. [doi: 10.1080/08952841.2014.888881] [Medline: 24919105]
71. Walker J, Crotty BH, O'Brien J, Dierks MM, Lipsitz L, Safran C. Addressing the challenges of aging: how elders and their care partners seek information. Gerontologist 2017 Oct;57(5):955-962. [doi: 10.1093/geront/gnw060] [Medline: 27053506]
72. Given LM, Ruecker S, Simpson H, Sadler E, Ruskin A. Inclusive interface design for seniors: Image-browsing for a health information context. J Am Soc Inf Sci 2007 Sep;58(11):1610-1617. [doi: 10.1002/asi.20645]
73. Marinescu V, Rodat S. Romanian and German seniors in quest of online health-related information: an exploratory comparative study. Ro J Comm & PR 2018 Apr;20(1):25-45. [doi: 10.21018/rjcpr.2018.1.251]
74. Turner AM, Osterhage KP, Taylor JO, Hartzler AL, Demiris G. A closer look at health information seeking by older adults and involved family and friends: design considerations for health information technologies. AMIA Annu Symp Proc 2018;2018:1036-1045 [FREE Full text] [Medline: 30815147]
75. Alpay LL, Toussaint PJ, Ezendam NPM, Rövekamp TAJM, Graafmans WC, Westendorp RGJ. Easing internet access of health information for elderly users. Health Informatics J 2016 Jul;10(3):185-194. [doi: 10.1177/1460458204045416]
76. Hardt JH, Hollis-Sawyer L. Older adults seeking healthcare information on the internet. Educational Gerontology 2007 Jun;33(7):561-572. [doi: 10.1080/03601270701364628]
77. Sanders K, Sánchez Valle M, Viñaras M, Llorente C. Do we trust and are we empowered by “Dr. Google”? Older Spaniards’ uses and views of digital healthcare communication. Public Relat Rev 2015 Dec;41(5):794-800. [doi: 10.1016/j.pubrev.2015.06.015]
78. Czaja SJ, Sharit J, Nair SN, Lee CC. Older adults and internet health information seeking. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 2009 Oct Presented at: 53rd Human Factors and Ergonomics Society Annual Meeting 2009; October 19-23, 2009; San Antonio p. 126-130. [doi: 10.1177/154193120905300204]
79. Campbell R. Consumer informatics: elderly persons and the internet. Perspect Health Inf Manag 2005 May;2:2 [FREE Full text] [Medline: 18066370]
80. Campbell RJ, Nolfi DA. Teaching elderly adults to use the Internet to access health care information: before-after study. J Med Internet Res 2005 Jun;7(2):e19 [FREE Full text] [doi: 10.2196/jmir.7.2.e19] [Medline: 15998610]
81. Hoffman-Goetz L, Friedman DB, Celestine A. Evaluation of a public library workshop. J Consum Health Internet 2006 Sep;10(3):29-43. [doi: 10.1300/j381v10n03_03]
82. Leung A, Ko P, Chan K, Chi I, Chow N. Searching health information via the web: Hong Kong Chinese older adults' experience. Public Health Nurs 2007 Mar;24(2):169-175. [doi: 10.1111/j.1525-1446.2007.00621.x] [Medline: 17319889]
83. Campbell RJ. Internet-based health information seeking among low-income, minority seniors living in urban residential centers. Home Health Care Manag Pract 2008 Jul;21(3):195-202. [doi: 10.1177/1084822308322648]
84. Xie B, Bugg JM. Public library computer training for older adults to access high-quality Internet health information. Libr Inf Sci Res 2009 Sep;31(3):155-162 [FREE Full text] [doi: 10.1016/j.lisr.2009.03.004] [Medline: 20161649]
85. Chu A, Mastel-Smith B. The outcomes of anxiety, confidence, and self-efficacy with internet health information retrieval in older adults: a pilot study. Comput Inform Nurs 2010 Jul;28(4):222-228. [doi: 10.1097/NCN.0b013e3181e1e271] [Medline: 20571374]
86. Fink A, Beck JC. Developing and evaluating a website to guide older adults in their health information searches: a mixed-methods approach. J Appl Gerontol 2015 Aug;34(5):633-651. [doi: 10.1177/0733464813486961] [Medline: 24652883]
87. Park E, Kwon M. Health-related internet use by children and adolescents: systematic review. J Med Internet Res 2018 Apr;20(4):e120 [FREE Full text] [doi: 10.2196/jmir.7731] [Medline: 29615385]
88. Freeman JL, Caldwell PH, Bennett PA, Scott KM. How adolescents search for and appraise online health information: a systematic review. J Pediatr 2018 Apr;195:244-255.e1. [doi: 10.1016/j.jpeds.2017.11.031] [Medline: 29398062]
89. Detlefsen EG. Where am I to go? Use of the internet for consumer health information by two vulnerable communities. Libr Trends 2004;53(2):283-300 [FREE Full text]
90. Feufel MA, Stahl SF. What do web-use skill differences imply for online health information searches? J Med Internet Res 2012 Jun;14(3):e87 [FREE Full text] [doi: 10.2196/jmir.2051] [Medline: 22695686]

91. Lee SB, Oh JH, Park JH, Choi SP, Wee JH. Differences in youngest-old, middle-old, and oldest-old patients who visit the emergency department. Clin Exp Emerg Med 2018 Dec;5(4):249-255 [FREE Full text] [doi: 10.15441/ceem.17.261] [Medline: 30571903]
92. Asla T, Williamson K, Mills J. The role of information in successful aging: the case for a research focus on the oldest old. Libr Inf Sci Res 2006 Mar;28(1):49-63. [doi: 10.1016/j.lisr.2005.11.005]
93. Heinström J, Sormunen E, Savolainen R, Ek S. Developing an empirical measure of everyday information mastering. J Assoc Inf Syst 2019 Oct;71(7):729-741. [doi: 10.1002/asi.24305]
94. Lee K, Hoti K, Hughes JD, Emmerton L. Dr Google and the consumer: a qualitative study exploring the navigational needs and online health information-seeking behaviors of consumers with chronic health conditions. J Med Internet Res 2014 Dec;16(12):e262 [FREE Full text] [doi: 10.2196/jmir.3706] [Medline: 25470306]
95. Marton C, Choo CW. A review of theoretical models of health information seeking on the web. J Doc 2012 Apr;68(3):330-352. [doi: 10.1108/00220411211225575]
96. Sun Y, Zhang Y, Gwizdka J, Trace CB. Consumer evaluation of the quality of online health information: systematic literature review of relevant criteria and indicators. J Med Internet Res 2019 May;21(5):e12522 [FREE Full text] [doi: 10.2196/12522] [Medline: 31045507]
97. Kong W, Song S, Zhao YC, Zhu Q, Sha L. TikTok as a health information source: assessment of the quality of information in diabetes-related videos. J Med Internet Res 2021 Sep;23(9):e30409 [FREE Full text] [doi: 10.2196/30409] [Medline: 34468327]
98. Song S, Zhao YC, Yao X, Ba Z, Zhu Q. Short video apps as a health information source: an investigation of affordances, user experience and users' intention to continue the use of TikTok. Internet Res 2021 Jun;31(6):2120-2142 [FREE Full text] [doi: 10.1108/intr-10-2020-0593]
99. Zhao YC, Xu X, Sun X, Zhu Q. An integrated framework of online generative capability: interview from digital immigrants. Aslib J Inf Manag 2014 Mar;66(2):219-239. [doi: 10.1108/ajim-08-2013-0068]
100. Zhao M, Song S, Zhao Y. Health information seeking on social apps among older adults living with chronic conditions. In: Proceedings of the Association for Information Science and Technology. 2021 Oct Presented at: Proceedings of the Association for Information Science and Technology; October 29-November 2, 2021; Salt Lake City, United States p. 878-880. [doi: 10.1002/pra2.595]
101. Manafo E, Wong S. Exploring older adults' health information seeking behaviors. J Nutr Educ Behav 2012 Jan;44(1):85-89. [doi: 10.1016/j.jneb.2011.05.018] [Medline: 22101129]
102. Song X, Song S, Chen S, Zhao YC, Zhu Q. Factors influencing proxy internet health information seeking among the elderly in rural China: a grounded theory study. In: Human Aspects of IT for the Aged Population. Design for the Elderly and Technology Acceptance. HCII 2019. Lecture Notes in Computer Science. 2019 Presented at: International Conference on Human-Computer Interaction; July 26-31, 2019; Orlando, United States p. 332-343. [doi: 10.1007/978-3-030-22012-9_24]
103. Arthanat S. Promoting information communication technology adoption and acceptance for aging-in-place: a randomized controlled trial. J Appl Gerontol 2021 May;40(5):471-480. [doi: 10.1177/0733464819891045] [Medline: 31782347]
104. Jung EH, Sundar SS. Status update: gratifications derived from Facebook affordances by older adults. New Media Soc 2018 Apr;20(11):4135-4154. [doi: 10.1177/1461444818768090]
105. Jung EH, Sundar SS. Senior citizens on Facebook: How do they interact and why? Comput Hum Behav 2016 Aug;61:27-35. [doi: 10.1016/j.chb.2016.02.080]
106. Zhao YC, Zhang Y, Tang J, Song S. Affordances for information practices: theorizing engagement among people, technology, and sociocultural environments. J Doc 2020 Oct;77(1):229-250. [doi: 10.1108/JD-05-2020-0078]
107. Rathnayake C, Winter JS. Carrying forward the uses and grats 2.0 Agenda: an affordance-driven measure of social media uses and gratifications. J Broadcast Electron Media 2018 Jul;62(3):371-389. [doi: 10.1080/08838151.2018.1451861]
108. Zheng Y, Hatakka M, Sahay S, Andersson A. Conceptualizing development in information and communication technology for development (ICT4D). Inf Technol Dev 2017 Nov;24(1):1-14. [doi: 10.1080/02681102.2017.1396020]
109. Hevner AR, March ST, Park J, Ram S. Design science in information systems research. MIS Q 2004 Mar;28(1):75-105. [doi: 10.2307/25148625]
110. Gregor S, Hevner AR. Positioning and presenting design science research for maximum impact. MIS Q 2013 Jun;37(2):337-355. [doi: 10.25300/misq/2013/37.2.01]
111. Webster J, Watson RT. Analyzing the past to prepare for the future: writing a literature review. MIS Q 2002 Jun;26(2):13-23. [doi: 10.2307/4132319]

Abbreviations

ICT: information and communication technology
OHIS: online health information seeking

Edited by T Leung; submitted 08.11.21; peer-reviewed by Y Li, A Alasmari, W Pian; comments to author 28.11.21; revised version received 18.12.21; accepted 12.01.22; published 16.02.22.

Please cite as:
Zhao YC, Zhao M, Song S
Online Health Information Seeking Behaviors Among Older Adults: Systematic Scoping Review
J Med Internet Res 2022;24(2):e34790
URL: https://www.jmir.org/2022/2/e34790
doi:10.2196/34790
PMID:35171099

Review

Sensing Apps and Public Data Sets for Digital Phenotyping of Mental Health: Systematic Review

Jean P M Mendes[1], BSc; Ivan R Moura[1], MSc; Pepijn Van de Ven[2], PhD; Davi Viana[1], PhD; Francisco J S Silva[1], PhD; Luciano R Coutinho[1], PhD; Silmar Teixeira[3], PhD; Joel J P C Rodrigues[4,5], PhD; Ariel Soares Teles[1,3,6], PhD

[1]Laboratory of Intelligent Distributed Systems, Federal University of Maranhão, São Luís, Brazil

[2]Health Research Institute, University of Limerick, Limerick, Ireland

[3]NeuroInovation & Technological Laboratory, Federal University of Delta do Parnaíba, Parnaíba, Brazil

[4]College of Computer Science and Technology, China University of Petroleum (East China), Qingdao, China

[5]Instituto de Telecomunicações, Covilhã, Portugal

[6]Federal Institute of Maranhão, Araioses, Brazil

Corresponding Author:
Ariel Soares Teles, PhD
Federal Institute of Maranhão
Rua José de Alencar, S/N, Bairro Cumprida
Araioses, 65570-000
Brazil
Phone: 55 86995501313
Email: ariel.teles@ifma.edu.br

Abstract

Background: Mental disorders are normally diagnosed exclusively on the basis of symptoms, which are identified from patients' interviews and self-reported experiences. To make mental health diagnoses and monitoring more objective, different solutions have been proposed such as digital phenotyping of mental health (DPMH), which can expand the ability to identify and monitor health conditions based on the interactions of people with digital technologies.

Objective: This article aims to identify and characterize the sensing applications and public data sets for DPMH from a technical perspective.

Methods: We performed a systematic review of scientific literature and data sets. We searched 8 digital libraries and 20 data set repositories to find results that met the selection criteria. We conducted a data extraction process from the selected articles and data sets. For this purpose, a form was designed to extract relevant information, thus enabling us to answer the research questions and identify open issues and research trends.

Results: A total of 31 sensing apps and 8 data sets were identified and reviewed. Sensing apps explore different context data sources (eg, positioning, inertial, ambient) to support DPMH studies. These apps are designed to analyze and process collected data to classify (n=11) and predict (n=6) mental states/disorders, and also to investigate existing correlations between context data and mental states/disorders (n=6). Moreover, general-purpose sensing apps are developed to focus only on contextual data collection (n=9). The reviewed data sets contain context data that model different aspects of human behavior, such as sociability, mood, physical activity, sleep, with some also being multimodal.

Conclusions: This systematic review provides in-depth analysis regarding solutions for DPMH. Results show growth in proposals for DPMH sensing apps in recent years, as opposed to a scarcity of public data sets. The review shows that there are features that can be measured on smart devices that can act as proxies for mental status and well-being; however, it should be noted that the combined evidence for high-quality features for mental states remains limited. DPMH presents a great perspective for future research, mainly to reach the needed maturity for applications in clinical settings.

(J Med Internet Res 2022;24(2):e28735) doi:10.2196/28735

KEYWORDS

mental health; digital phenotyping; sensing apps; data sets; sensor data

Introduction

Background

Mental health issues have a high prevalence, with 1 in 10 people worldwide experiencing them at any one time [1] and common mental disorders such as depression being closely linked to suicide [2]. Mental disorders are "generally characterized by some combination of abnormal thoughts, emotions, behavior and relationships with others" [3]. Examples are depression, schizophrenia, excessive anxiety and stress, disorders caused by drug and alcohol abuse, and personality and delusional disorders. These disorders pose a significant burden on societies, both emotionally and financially. For example, the cost of mental health disorders in the European Union is estimated at €600 billion (~US $451 billion), or 4% of gross domestic product [4]. COVID-19 has had a further negative impact on global mental health [5].

Mental disorders are usually diagnosed exclusively on the basis of symptoms, which are identified from patients' interviews and self-reported experiences. Sometimes these experiences are gathered using ecological momentary assessment (EMA) solutions [6], but mostly therapists rely on patients remembering such experiences during sessions. EMA solutions are used as a research method to collect, at fixed or random moments, reports from individuals about perceptions of their behaviors and feelings, and what they have done or experienced. It is well known that the intervening time and current state of the patient bias his/her memory of the experience. In addition, biological tests to assist diagnosis remain hard to be developed [7]. Based on the need to develop solutions able to objectively diagnose and monitor mental health, different solutions have been proposed, such as mobile apps [8,9] and machine learning (ML) solutions [10], which are even more indicated today due to the global pandemic situation [11,12]. Digital phenotype solutions are examples that can expand the ability to identify and diagnose health conditions from the interactions of people with digital technologies [13]. Specifically, digital phenotyping of mental health (DPMH) [14] seems to be a promising approach not only to deal with the problem of diagnosing the issue, but also to be applied to the treatment.

The omnipresent adoption of pervasive devices, including smartphones and wearable sensors, provides novel opportunities for tracking mental health status and disorders. Digital phenotyping refers to the "moment-by-moment quantification of the individual-level human phenotype in-situ using data from smartphones and other personal digital devices" [15], thereby removing limitations created by the aforementioned bias in self-reports.

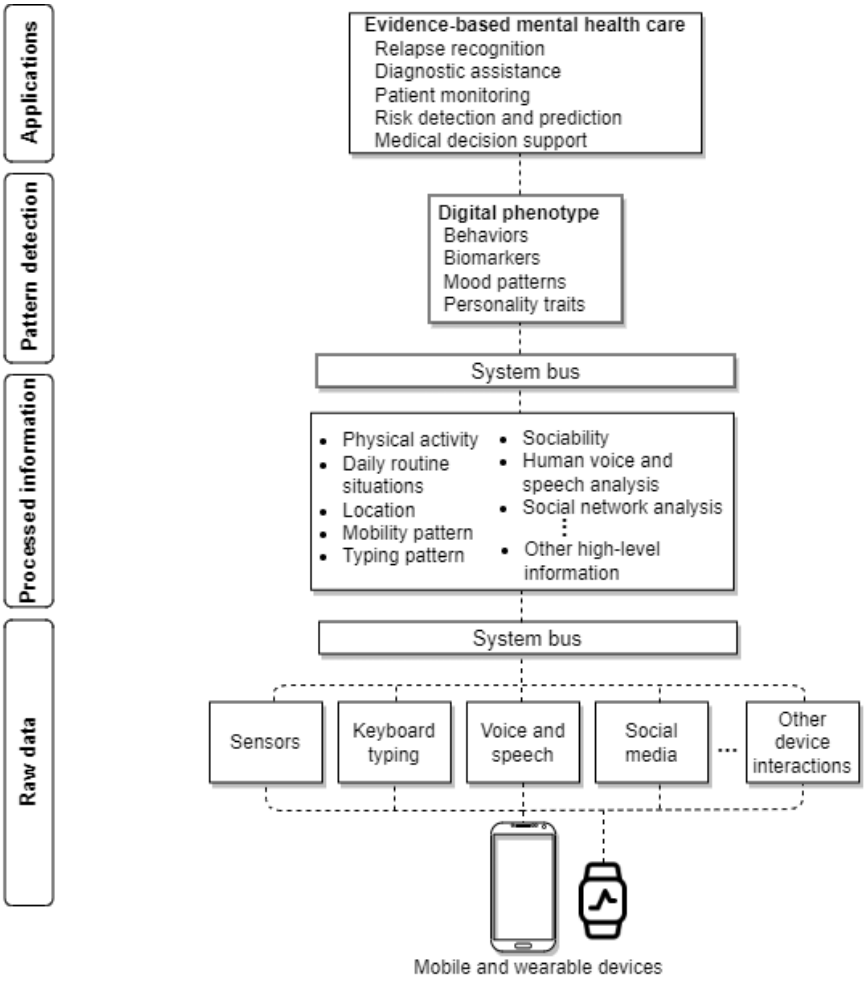
DPMH solutions require collecting and analyzing large amounts of different types of social and behavioral data that can represent experiences of the users and their interactions with people, places, and devices. These context data can be passively gathered, for instance, from ubiquitous sensors, social media, and health care systems [16]. After collection, pieces of raw data are usually preprocessed and transformed into useful data or data sets to be mined [17]. For example, these data sets may be analyzed or used as input to build ML models [18], including for DPMH, to produce valuable insights and evidence. Therefore, DPMH sensing apps are primarily responsible for collecting and preprocessing data, with the data sets produced being important for developing such models. This study systematically reviews the sensing apps and data sets for DPMH.

Definitions

In the last few years, the number of smart devices, that is, mobile (eg, smartphone, tablet) and wearable (eg, smart band, smartwatch) devices, has grown globally. They have enabled the development of research in the health area, including mental health [10]. The term "digital phenotype", defined by Jain and colleagues [13], refers to the identification of human behavior patterns, whereas "digital phenotyping" is a monitoring approach that can collect patients' behavioral markers passively [19]. Therefore, DPMH solutions aim at collecting multimodal pieces of information from digital devices using sensing apps to combine them with electronic medical records to objectively contribute to the identification of symptoms of mental disorders. In this context, sensing apps are tools for mobile and wearable devices used to collect useful user information.

Our vision of the digital phenotyping process organized in layers is presented in Figure 1. The process starts at the first layer with the collection of raw data from different sources (eg, global positioning system [GPS] sensors, keyboard inputs, voice, and social media). These data can be collected both actively, in which user inputs are explicitly required, and passively [20], which only requires the user's permission to access context data. In the next layer, these data are processed to provide high-level information. High-level information represents not only human behaviors (eg, sociability, physical activity) and habits (eg, mobility, sleep) but also other information of interest for professionals (eg, environmental context, mood). Next, human behavioral patterns that compose digital phenotypes (eg, biomarkers, mood patterns) can be recognized using computational tools (eg, ML, data mining, statistical models). Finally, we visualize the application layer, which corresponds to digital phenotypes used by health professionals for evidence-based mental health care.

Figure 1. The process of digital phenotyping.



Related Work

Since the aforementioned concepts were proposed in the literature, many research studies have been performed. For this reason, researchers have also reviewed different aspects regarding this research topic. Table 1 presents a list composed of related reviews.

Table 1. List of related review articles.

Study	Description
Garcia-Ceja et al [10]	A survey on mental health monitoring using mobile and wearable sensors focused on multimodal sensing and machine learning solutions.
Cornet and Holden [21]	An SLR[a] on passive sensing using specifically smartphones focused on health and well-being.
De-La-Hoz-Franco et al [22]	An SLR aimed at finding data sets composed of sensor data for human activity recognition.
Trifan et al [23]	This SLR aimed to identify studies on the passive use of smartphones for generating outcomes related to health and well-being. It identified that one of the areas most explored by mobile passive sensing is mental health.
Seppälä et al [24]	An SLR on mobile solutions focused on uncovering associations between sensor data and symptoms of mental disorders (ie, behavioral markers).
Liang et al [14]	A comprehensive survey addressing different topics on DPMH[b].
Benoit et al [20]	This SLR sought to map DPMH tools that use machine learning algorithms across the schizophrenia spectrum and bipolar disorders.
Antosik-Wójcińska et al [25]	This work presents an overview of studies about smartphone systems focused on monitoring or detecting bipolar disorder.

[a]SLR: systematic literature review.

[b]DPMH: digital phenotyping of mental health.

This review differs from the previous ones in the following aspects: First, instead of focusing on a specific mental state/disorder, this review presents an overview of how different types of devices and detection modalities have been used to monitor a wide variety of different mental states within the DPMH area. Second, this review covers not only active collection solutions, which are emphasized in most reviews, but also passive sensing proposals. Third, this review focuses on the technical features of sensing apps and data sets (eg, size, sensors used to collect data, and types of context data). Technical features can be identified to serve as a basis for the use or development of new apps (eg, physical and virtual sensors used to collect data, operating systems for which the apps were developed, types of context data collected, inferred information). Finally, not all previous reviews were conducted systematically. Our article therefore provides researchers with an overview of the available technological framework for DPMH and can serve as a preliminary guide for current and further research.

Objectives and Research Questions

This systematic review intends to provide a technical characterization and summary of sensing apps and public data sets for DPMH. By "public" we mean data sets that are available for free download for use in other research endeavors. These 2 topics (ie, sensing apps and public data sets) are jointly addressed in this review as complementary content. When researchers do not have access to DPMH data sets, they need sensing apps. This paper therefore can be a starting point not only to gain knowledge on the current sensing apps for DPMH (which consequently enables the development of new solutions), but also to find reusable ones. Therefore, the objectives of this article are to (1) present results from a systematic search on digital libraries and data set repositories, and then identify and categorize them by considering their characteristics; (2) summarize their main features (measurable pieces of data that can be used for analysis or creation of ML models, such as data collection time stamp, context data produced by DPMH solutions, and data self-reported by users), which are useful for researchers, either mental health or information technology ones, to conduct further investigation and comment on their usefulness; and (3) identify trends in and research opportunities for DPMH. Results of this systematic review are also relevant for data engineers and ML specialists who make efforts in developing DPMH solutions.

To achieve the objectives of this systematic review, we defined the following research questions for sensing apps (SA-RQs) and data sets (DS-RQs):

SA-RQ1: What context data are collected through DPMH sensing apps?

SA-RQ2: What high-level information can be inferred from the context data collected by DPMH sensing apps?

SA-RQ3: How is the identified high-level information used to support mental health?

DS-RQ1: What features are available in public data sets for DPMH?

DS-RQ2: What high-level information can be derived from public data sets for DPMH?

Methods

Design

This study was conducted based on the guidelines for systematic literature reviews in software engineering proposed by Kitchenham and Charters [26]. This review followed 3 main phases: planning, conducting research, and dissemination of results. These phases were supported by the *Parsif.al* [27] tool, which provides an online shared work environment for planning and executing systematic reviews. In this section, we present how this review was planned and conducted.

Search Strategy

The search aimed to identify data sets and studies that have presented sensing apps capable of collecting data. Two (JM and IM) researchers conducted an exhaustive search on January 14, 2021, on data set repositories and digital libraries. The search for data sets was performed in 20 repositories (Multimedia Appendix 1). The search for articles reporting sensing apps was conducted in the following digital libraries: ACM Digital Library, DOAJ, IEEE Xplore, Web of Science, PubMed, PsycInfo, ScienceDirect, and Scopus. These databases were selected because they collect reliable studies related to mental health informatics.

We designed the search strings to retrieve data sets and articles presenting sensing apps for DPMH (Table 2). These search strings were carefully designed to meet the research focus. In the string to search data sets, we defined the 2 main terms (ie, mental health and digital phenotyping) and decided to use Boolean "OR" as the link for them to get comprehensive results. The search string for articles was developed based on the review objective, research questions, and their motivations. We used keywords and their synonyms to maximize results. To avoid missing papers, we evaluated the suitability of the string in a pilot search, in which we used those studies developed by Liang et al [14] (ScienceDirect) and Torous et al [15] (PubMed) as control articles. This pilot search was able to retrieve the cited studies, thus demonstrating its ability to find articles relevant for this review. At the end of the search, duplicate data sets and articles were identified and removed using the *Parsif.al* tool.

Table 2. Keywords and their synonyms.

Search	Source	String
Data sets	Data set repositories	"mental health" OR "digital phenotyping"
Sensing apps	Digital libraries	("mental health" OR "mental disorder*" OR "mental illness" OR "mental state" OR "mental disease") AND ("mobile device" OR "smartphone*" OR "wearable device*" OR "sensor*" OR "wearable*" OR "mobile application*" OR "mobile health" OR "mHealth" OR "mobile phone*" OR "sensor data") AND ("passive detection" OR "data collection" OR "digital phenotype" OR "digital phenotyping" OR "digital health" OR "monitoring" OR "passive sensing")

Selection Criteria

A set of selection criteria was defined to track research articles and data sets. Textbox 1 presents the selection criteria for scientific studies with sensing apps and data sets. Importantly, no date range limits were applied to the literature included in the review. In the selection of scientific articles, criterion EC1 excluded studies presenting the development of EMA apps, and papers that do not present a new DPMH solution (eg, studies using a DPMH solution previously described/published in another paper). For data set selection, criterion EC1 excluded those data sets that were not publicly available, that is, those protected and not accessible to be reused by other researchers.

In the selection phase, 2 researchers (JM and IM) performed the data set selection process based on the inclusion and exclusion criteria. In a second step, the same 2 researchers independently performed the study selection process. This process consisted of 3 sequential phases: (1) study screening by means of metadata analysis (ie, title, abstract, and keywords); (2) full-text analysis of the articles selected in the screening phase; and (3) conducting backward snowballing [28]. Next, the level of agreement between the selections was calculated using the Cohen κ coefficient [29]. In the end, the 2 researchers conducted discussions to resolve selection conflicts and, when there was no consensus, judges (2 other authors, namely, AT and DV) deliberated on the disagreements.

Textbox 1. Selection criteria.

Inclusion criteria (IC)

Scientific articles

IC1: Primary studies that present pervasive solutions to collect data for digital phenotyping of mental health.

IC2: Full papers.

IC3: Papers in English language.

Data sets

IC1: Available to be downloaded and used in other research studies (ie, public data set).

IC2: Focused on mental health or specific mental disorders.

IC3: Relevant data (eg, behavioral, physiological, social) for mental health collected through pervasive technologies.

IC4: Content in English language.

Exclusion criteria (EC)

Scientific articles

EC1: Articles presenting research on digital phenotyping of mental health without involving a proposal of a pervasive solution.

EC2: Gray literature.

EC3: Articles that have other publications with a more current and complete version of the proposed solution.

Data sets

EC1: Not publicly available.

EC2: With no content related to mental health.

EC3: Data on treatments of patients with mental disorders without using pervasive devices.

EC4: Content in languages different from English.

EC5: Online surveys on ethnographic characteristics and prevalence of mental disorders.

EC6: Composed exclusively by multimedia data (eg, video, audio) or electroencephalography data.

Data Extraction

In this step, data were extracted from the selected articles and data sets to answer the research questions defined in this review. For this purpose, a data extraction form was designed by 2 authors (JM and IM) and validated by the judges. Specifically, we designed the items in the form to extract relevant information presented by the reviewed studies and data sets, thus enabling

us to answer the research questions, and identify open issues and research trends. Multimedia Appendix 2 presents the items in the data extraction form.

Results

Study Selection

An overview of the review process with results is presented in Figures 1 and 2. In Figure 2, 8 digital libraries were used to search for scientific articles that presented sensing apps for DPMH. A total of 2374 articles were returned. We removed 926 duplicate articles. The inclusion and exclusion criteria from Textbox 1 were applied to select 26 selected studies. The Cohen κ statistical test showed an agreement level of ≈0.87 between researchers, which is considered an almost perfect agreement [29]. Next, researchers used the 1-level backward snowballing approach and added 5 articles. This resulted in 31 articles for inclusion in the data extraction process.

In Figure 3, 20 data set repositories were searched to return 2581 data sets with 471 duplicates that were removed. After applying selection criteria (Textbox 1) and resolving conflicts, 8 data sets remained for analysis.

Figure 2. PRISMA-based flowchart describing the selection of studies.

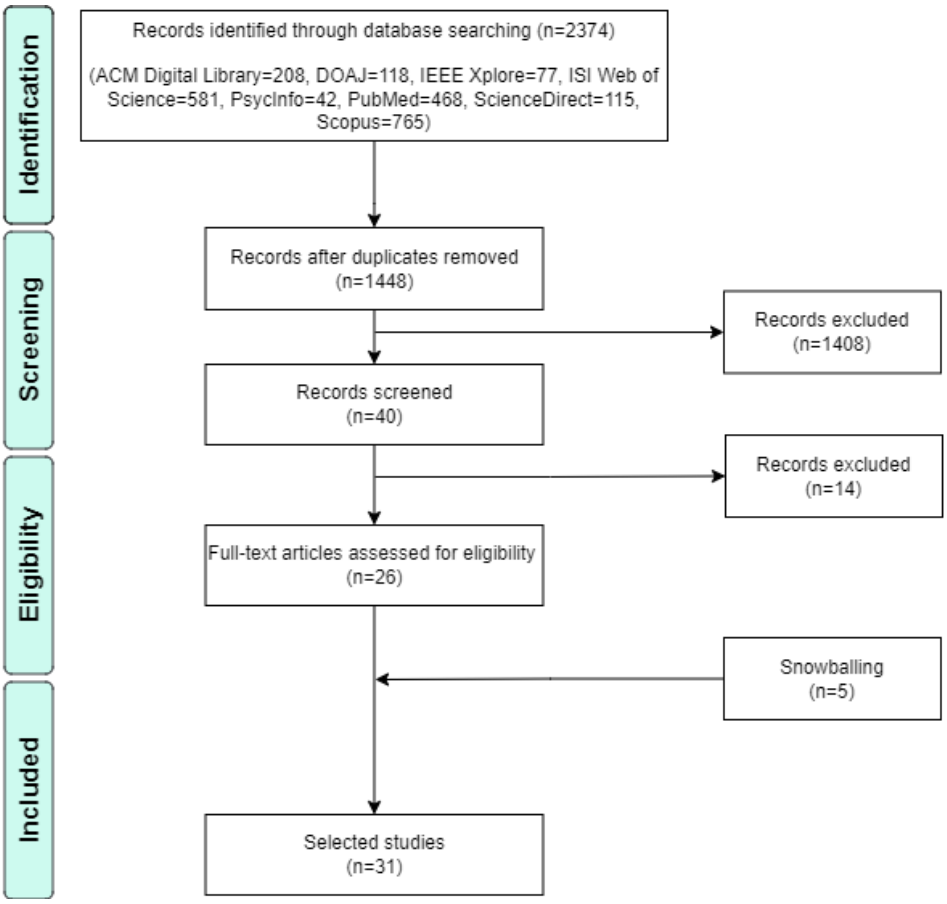
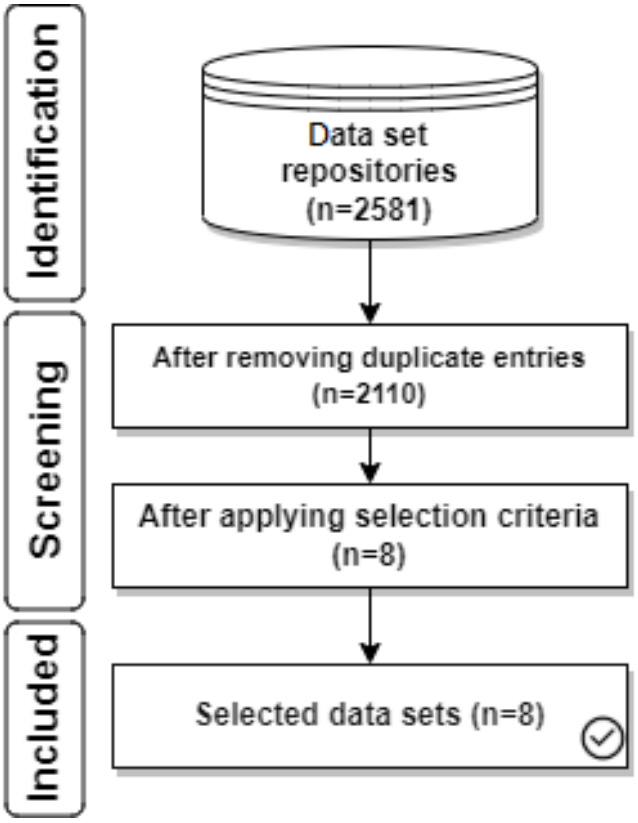


Figure 3. Flowchart describing the selection of data sets.



Sensing Apps

Table 3 summarizes the 31 apps identified, which are presented in ascending order by year of publication. Multimedia Appendix 3 presents the full version of the table. Context data sources are categorized as follows to present the sensors used by the apps based on the work by Palaghias et al [30]: ambient (eg, microphone, camera), positioning (eg, GPS, Wi-Fi), virtual (eg, phone calls, SMS text messages), and inertial (eg, accelerometer, gyroscope). Table 3 also presents high-level information inferred and types of analyses performed on the collected data. Apps that do not infer information (ie, defined as "It does not infer information") are only intended to collect data from smart devices. In this case, collected data are usually sent to servers for analysis. These apps are flagged as "Raw data collection" in Table 3.

Table 3. Summary of reviewed sensing apps.

App	Context data source	High-level information	Type of analysis
Funf [31]	Positioning, inertial, and virtual	It does not infer information	Raw data collection
Mobilyze [32]	Positioning, inertial, virtual, and ambient	Mood, emotions, cognitive/motivational states, physical activity, social context	Mental state prediction
Purple Robot [33]	Positioning, inertial, and virtual	It does not infer information	Raw data collection
AWARE [34]	Positioning, inertial, and virtual	It does not infer information	Raw data collection
Sensus [35]	Positioning, inertial, virtual, and ambient	It does not infer information	Raw data collection
MOSS [36]	Positioning and virtual	Physical activity, mobility, device usage, sociability, app usage	Mental state classification
Beiwe [15]	Positioning, inertial, virtual, and ambient	It does not infer information	Raw data collection
EVO [37]	Positioning, inertial, and virtual	It does not infer information	Raw data collection
CrossCheck [38]	Positioning, inertial, virtual, and ambient	Sleep, sociability, mobility, physical activity, device usage	Mental state prediction
SituMan [39]	Positioning and inertial	Daily routine situations (eg, working, studying)	It recognizes daily routine situations using fuzzy logic
EmotionSense [40]	Positioning, inertial, virtual, and ambient	Semantic locations, physical activity, sociability	Correlation analysis and mental state classification
StudentLife [41]	Positioning, inertial, virtual, and ambient	Sociability, mobility, physical activity, device usage	Correlation analysis
Undefined [42]	Positioning, inertial, and ambient	Physical activity, mobility, and sociability	Correlation analysis
AMoSS [43]	Positioning	Mobility	Mental state prediction
eB2 [44]	Positioning and virtual	Mobility	Mental state classification
EARS [45]	Positioning, inertial, virtual, and ambient	It does not infer information	Raw data collection
SleepGuard [46]	Inertial and ambient	Posture/position of body when sleeping	Mental state classification
Moment [47]	Virtual	It does not infer information	Mental state classification
TypeOfMood [48]	Virtual	It does not infer information	Mental state classification
RADAR-base [49]	Positioning, inertial, virtual, and ambient	It does not infer information	Raw data collection
SHADO [50]	Positioning, inertial, and ambient	Physical activity, mobility, sleep, sociability	Correlation analysis and mental state classification
InSTIL [51]	Positioning, inertial, virtual, and ambient	It does not infer information	Raw data collection
Lamp [52]	Positioning	Physical activity	Correlation analysis
SOLVD [53]	Positioning, inertial, virtual and ambient	Mobility, sociability, context of daily life (eg, duration of sleep)	Correlation analysis
STDD [54]	Inertial, virtual, and ambient	Physical activity, mood, sociability, sleep	Mental state classification
Moodable [55]	Positioning, virtual, and ambient	Sociability and mobility	Mental state classification
Cogito Companion [56]	Positioning and Virtual	Mood, stress level, and well-being	Mental state classification
Strength Within Me [57]	Virtual	Sleep, mobility, and sociability	Mental state prediction
EuStress [58]	Ambient	It does not infer information	Mental state prediction
Mood Triggers [59]	Positioning, inertial, virtual, and ambient	Mobility and sociability	Mental state prediction
Data Collector [60]	Positioning and inertial	Physical activity and mobility	Mental state classification

Data Set Characterization

Table 4 shows the 8 selected data sets in descending order by number of participants. Two of them have sleep quality data: data sets DS1 and DS7, in which the data are derived from activity trackers such as Fitbit, smartwatches, and smartphones. We identified 2 data sets (DS3 and DS5) with data collected from various sensors, which we refer to as multimodal. We identified 2 data sets (DS3 and DS5) that were generated by the StudentLife [41] and Beiwe [15] sensing apps, respectively, shown in Table 3.

Table 4. Summary of DPMH data sets.

Data set	Study	High-level information	Features	Device type/operating System	Number of participants	Study duration	Size
DS1[a] [61]	[62]	Sleep quality	Fitbit data (eg, heart rate, sleep duration, sleep time, wake time)	Watch Fitbit	482	3-11 nights	392.32 KB
DS2 [63]	[64]	Activity	Actigraph (time stamp, activity measurement from the actigraph watch)	Actigraph watch	55	Average 12.6 days	4.3 MB
DS3 [65,66]	[41,67]	Multimodal (stress, sleep, mood, physical activity, sociability, well-being)	Self-report questionnaires, activity, audio, Bluetooth encounters, conversation, lightness, GPS[b] coordinates, phone charge, screen on/off, Wi-Fi IDs	Smartphone (Android)	48	66 days	230 MB/5 GB
DS4 [68]	[69,70]	Sociability	Self-reports, battery level, Bluetooth encounters	Smartphone (Android, iOS)	32	4 weeks	9.7 MB
DS5 [71]	[15]	Multimodal (mobility, sociability, sleep)	Self-report questionnaires, accelerometer, app logs, Bluetooth encounters, call logs, GPS coordinates, power state, Wi-Fi	Smartphone (Android, iOS)	6	3 months	776.7 MB
DS6 [72]	[73]	Mood, depression symptoms	Self-report questionnaires	Smartphone (Android, iOS)	3	14 days	2.7 MB
DS7 [74]	—	Sleep quality	Start, end, sleep quality, time in bed, wake-up time, sleep notes, heart rate, number of steps	Wearable device and smartphone (iOS)	1	4 years	66.11 KB
DS8 [75]	—	Mood	Self-reported mood	Mobile social network (Twitter app)	1	2 years	131 KB

[a]DS: data set.

[b]GPS: global positioning system.

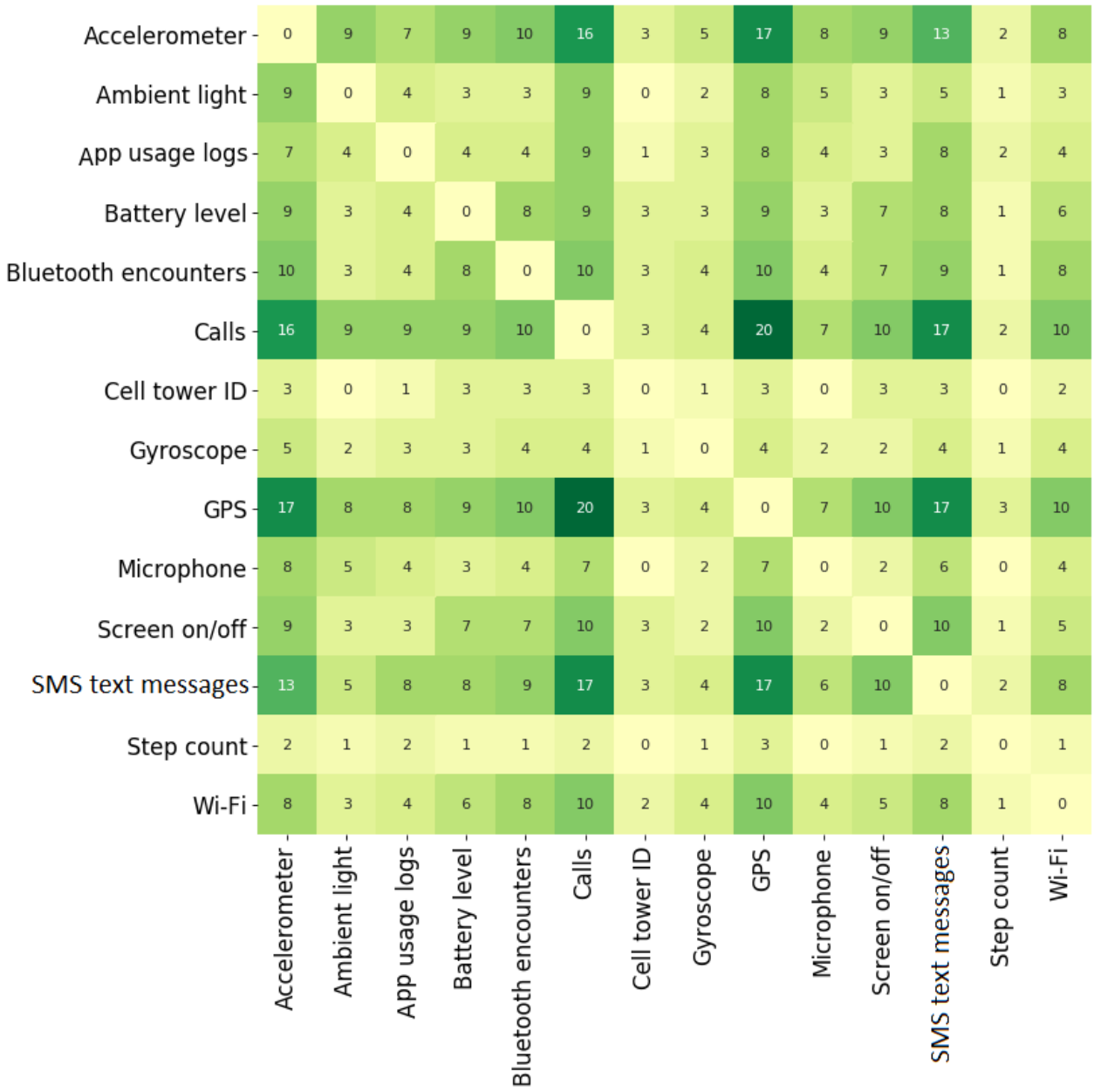
Data set DS1 [61] presents sleep data (eg, total sleep time and sleep efficiency) obtained from Fitbit Charge HR activity trackers used by 482 individuals [62], while data set DS2 [63] includes actigraphic data collected from patients with unipolar and bipolar disorders and 32 healthy controls [64]. Data set DS3 [65,66] contains data gathered from different sensors and EMA questionnaires collected from smartphones of 48 undergraduate and graduate students over 66 days [41,67]. Data set DS4 [68] comprises Bluetooth device scan, battery level, and EMA data collected at regular intervals for 4 weeks [69,70], while data set DS5 [71] presents passive data (eg, GPS, Wi-Fi, Bluetooth, and accelerometer) and active data (EMA survey responses) collected over 3 months [15]. Data set DS6 [72] contains EMA assessments of depression symptoms using the 9-item Patient Health Questionnaire (PHQ-9) [73]. Data set DS7 [74] presents sleep data collected through the Sleep Cycle mobile app [76]. Finally, data set DS8 [75] presents values extracted from Twitter posts collected from a person using Exist [77] over 2 years.

Context Data Collected by DPMH Sensing Apps (SA-RQ1)

Sensing apps identified in this review collect context data from mobile and wearable devices to support DPMH. At a high level, the sensors that measure context data can be seen as physical and virtual sensors [78], which generate a diversified set of behavioral data. Physical sensors are hardware components embedded or connected to devices responsible for collecting context data. Some examples are accelerometers to measure user activity, light sensors to measure ambient light levels, and GPS to collect user's locations. Virtual sensors represent software components capable of recording interactions of individuals with devices or using a number of physical sensors (or other virtual ones) to construct a higher-level feature. Examples of such sensors are social interaction sensors that may use Bluetooth encounters (ie, co-location information between individuals or places), Wi-Fi network, and sound data to infer social activity; and user–device interaction sensor, which measures user interactions with devices (eg, call logs, SMS text messages, app usage, screen on/off).

Figure 4 presents a heat map of the combination of context data sources for the 31 sensing apps, showing the most used sensors in DPMH solutions. In this analysis, we investigated the frequency of the combination of each type of context data source, highlighting the main sets of sensors explored by the sensing apps. For example, Bluetooth encounters are often combined with accelerometer (n=10), battery level (n=8), calls (n=10), GPS (n=10), screen on/off (n=7), SMS text messages (n=9), and Wi-Fi (n=8), while app usage logs are often combined with accelerometer (n=7), calls (n=9), GPS (n=8), and SMS text messages (n=8). We also identified from this analysis that step count (Fitbit), cell tower ID, and gyroscope are combined less often with other context data sources. The analysis of the combination of context data sources (Figure 4) demonstrates an interest in performing data fusion to identify multiple high-level information and emphasizes the combination of context data sources resulting from the interest in monitoring such information. For example, we identify an interest in recognizing sociability information by combining call logs with Bluetooth encounters (n=10) and SMS text messages (n=17). We also recognize that GPS is often combined with Wi-Fi (n=10) to recognize mobility aspects. In addition, the interest in monitoring multiple high-level information in the same app resulted in different combinations of context data sources. For example, the combination of GPS with call logs (n=20), accelerometer (n=17), and screen on/off (n=10) is a result of an interest in monitoring sociability, physical activity, and device usage patterns, respectively.

Figure 4. Context data sources used in the reviewed studies. GPS: global positioning system.



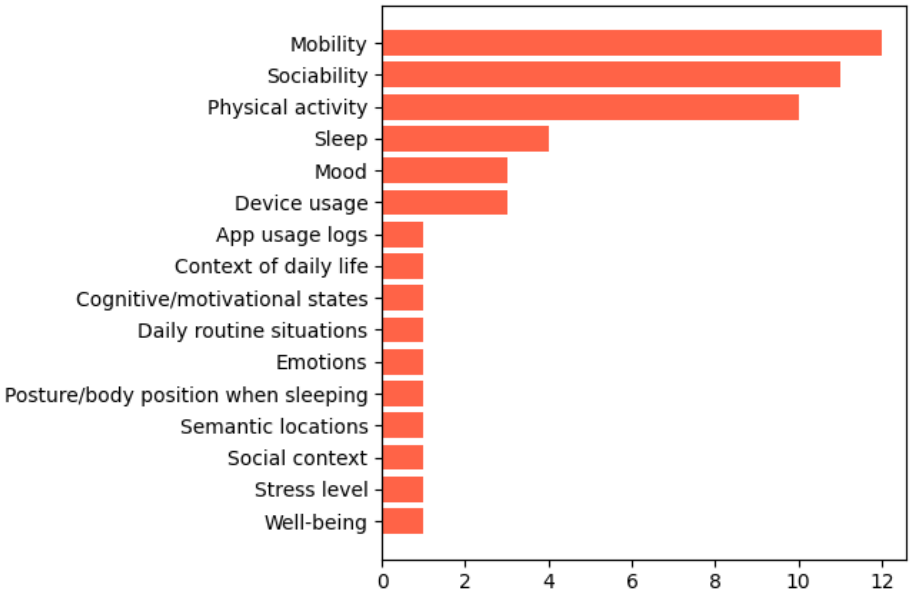
High-Level Information Identified by Sensing Apps (SA-RQ2)

From the context data collected by sensing apps, researchers can extract high-level information representing different types of situations (eg, sociability, mobility). Table 3 presents the situations of interest identified from context data. Sensing apps aimed to identify information related to the physical and environmental aspects of the monitored individuals, such as mobility patterns [38] (eg, places visited, total distance traveled, time spent in locations), physical activities (activity type and duration), daily routine situations (eg, working, studying), and environmental context (eg, ambient temperature).

Figure 5 shows the types of high-level information generated by the sensing apps. The 3 types of information that stand out are human behavioral patterns related to mobility, sociability, and physical activity (n≥10). Information about the individual's condition was also derived, such as mood and sleep quality.

Researchers also explored information about device usage, which was derived from logs such as calls, SMS text messages, screen on/off events, and app usage. In general, studies have been able to build apps that achieve promising results of performance metrics (eg, accuracy, sensitivity, specificity) in identifying useful high-level information for mental health professionals. By contrast, there are some researchers developing apps that have not transformed context data into high-level information (ie, they focus only on raw data collection), and these are not shown in Figure 5.

Figure 5. High-level information summary.

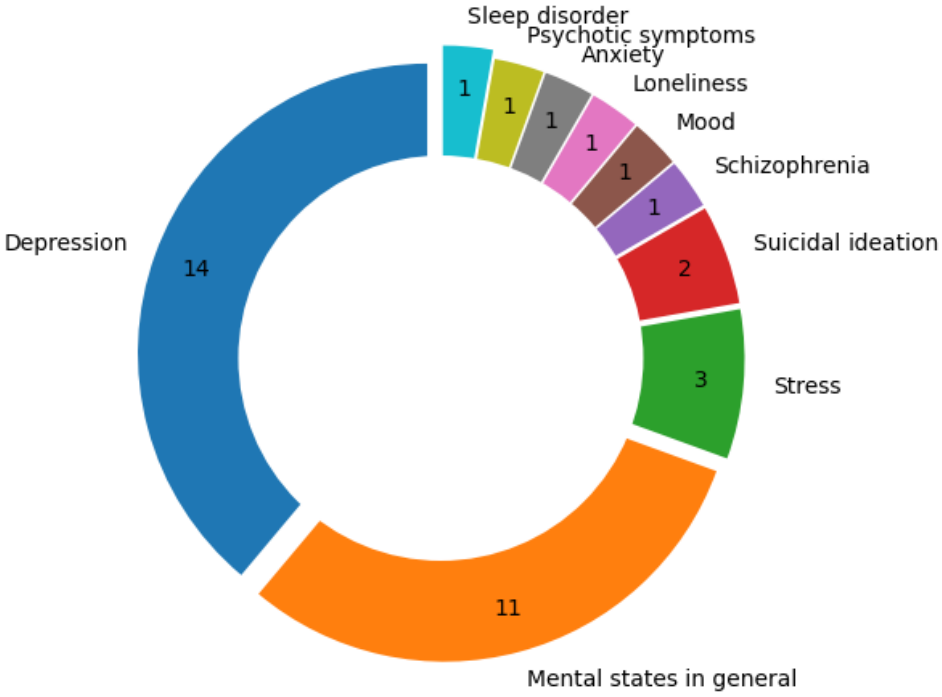


Support for Monitoring Mental Health (SA-RQ3)

The sensing apps identified used high-level information to provide a variety of mental health services. Table 3 shows the types of analyses performed based on the high-level information identified. Some apps infer daily routine situations and send recommendations in real time [58], thus aiming to provide tools to improve services of health professionals. Most approaches to support mental health monitoring were as follows: correlation, classification, and prediction. Correlation analyses associate features extracted from high-level information with mental states of the monitored individual, that is, they aim to find evidence that identified behaviors have significant correlations with psychological well-being [79]. Researchers also used identified behaviors to design ML models capable of classifying and predicting mental states [32,80], which can be used as decision support tools for health professionals. Lastly, some studies [31,81,82] did not report on additional analyses, but concentrated on describing the features of their sensing apps to facilitate DPMH research.

Figure 6 shows the mental states/disorders studied by DPMH research. Apps classified as "Mental states in general" did not focus on a specific mental disorder; instead, they are generic to be used in studies for different mental health disorders. We found 14 articles with a focus on individuals with depression. Other mental states/disorders are schizophrenia, mood, suicidal ideation, stress, loneliness, anxiety, and psychotic symptoms, all with between 1 and 3 studies returned in our search. We identified 11 articles that did not specifically address a particular mental disorder in their studies.

Figure 6. Mental states/disorders targeted by sensing apps.



Features Available in Data Sets (DS-RQ1)

The selected data sets have several types of features extracted from context data collected by sensing apps. These features model various aspects of human behavior that can be applied to the development process of new tools for monitoring and intervention in mental health. Table 4 presents the features available in the selected data sets. Data sets DS1 and DS7 contain features related to sleep. They provide information such as sleep start and end, sleep quality, time in bed, wake-up, sleep notes. Data sets DS4 and DS8 have features related to the social aspect such as self-reports of social interactions and Bluetooth encounter data, while data sets DS6 and DS8 provide actigraph data and self-reports, respectively. Data sets DS4 and DS5 have features capable of modeling more than 1 human behavior (ie, multimodal), thus providing data from different sources. These sources provide multimodal context data that can be fused to generate meaningful high-level information [10]. Moreover, multimodal data sets can support DPMH research under different aspects of interest for professionals, such as patient's mobility and sociability.

Possible High-Level Information Derived From Data Sets (DS-RQ2)

The selected data sets have features capable of modeling different types of human behavior. Therefore, to understand the potential for applying these data to DPMH, we identified high-level information that can be derived from these data sets based on the available context data. Table 4 presents high-level information inferred. Explicitly, these data sets can model the situations listed in Textbox 2.

Additionally, some data sets contain high-level information such as mood, mental status, and mental disorder symptoms. These types of information are self-reported by participants using questionnaires (eg, PHQ-9) and EMA solutions through smart devices.

Textbox 2. Situations modeled by data sets.

Sociability

This can be quantified using context data that allow characterizing social relationships of the participants such as interactions on online social networks, and face-to-face and device-mediated interactions [83]. These data sets contain context data such as posts on social networks, Bluetooth encounters, global positioning system (GPS) coordinates, or conversational activity inferred from microphone signals.

Physical activity

This is routinely measured using accelerometer and GPS data, resulting in either a log of user physical activities or an aggregate measure of energy expenditure.

Sleep

This is mostly measured in terms of sleep quality and sleep duration of the participants. In general, these data sets have features such as sleep quality, total sleep time, time in bed, and wake-up inferred from contextual data such as heart rate and screen on/off logs, and ambient light.

Multimodal

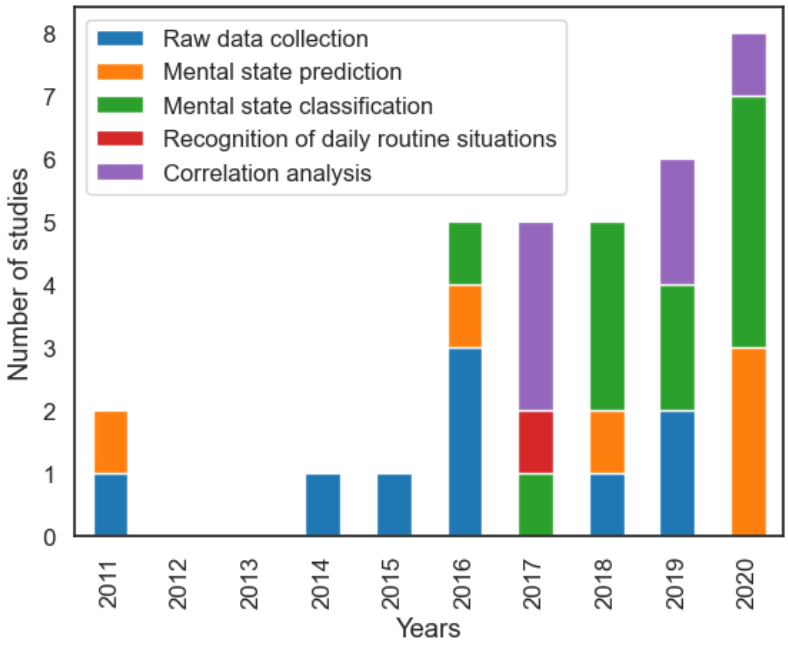
These data sets comprise several types of context data (eg, accelerometer, ambient light, battery level, Bluetooth, GPS, screen on/off, questionnaires [9-item Patient Health Questionnaire]), which allow characterizing more than 1 behavior of the participants such as sociability, mobility, and physical activity.

Discussion

Principal Findings

Our review shows that there are features that can be measured on smart devices that can act as proxies for mental status and well-being, but it should be noted that the combined evidence for high-quality features for mental states remains limited. Researchers have conducted several types of analysis on the data collected. In principle, we recognize a trend to design features from the data collected (Figure 7) to train ML models capable of classifying mental states/disorders (n=11) and predicting future mental states/disorders (n=6). We also note a substantial effort in analyzing correlations between features designed from the collected data and mental states/disorders (n=6). This type of analysis aims to find evidence of the viability and usefulness of DPMH for clinical practice. Furthermore, there are apps that only collect raw context data (n=9) to be analyzed subsequently, and 1 app (SituMan [39]) focused on the recognition of daily routine situations.

Figure 7. Number of published studies by year and types of analysis.



The literature mostly reports on the measurement of mobility, sociability, sleep, physical activity, and mood. Mobility represents high-level information derived from the movement sequence of individuals. These patterns are identified by processing GPS and Wi-Fi samples, which allow for the recognition of mobility traces. Sociability is measured using context data sources such as call logs, SMS text messages, Bluetooth encounters, and microphone data. These pieces of data allow identifying physical and virtual social interactions. Sleep information is measured by contextual data fusion such as ambient light, movement activity, screen on/off, and ambient sound. In addition, researchers have used Fitbit data to recognize sleep quality. Physical activity is recognized using data from inertial sensors (eg, accelerometer, magnetometer, gyroscope), making it possible to classify different types of activities such as walking, running, and stationary. Finally, mood has been

recognized using different context data sources, such as accelerometer and heart rate monitor of wearable devices, combined with self-reports.

The different ways in which these features are inferred and reported make it impossible to compare results across studies, or combine data sets to achieve greater statistical power. For this reason, we believe the research community would benefit from a clear standard on the measurement of these behaviors. The data sets identified and studies in this review provide an interesting starting point for such consensus building. Particularly, the StudentLife data set [41] has been explored by many studies that propose solutions capable of supporting mental health professionals. Different solutions have used this data set to detect human behavioral patterns and perform association, classification, and prediction of mental states. For example, by using the StudentLife data set, Saeb et al [80] analyzed the correlation between mobility patterns identified from GPS samples and depressive symptoms reported by students. Farhan et al [84] designed a multiview biclustering model using various features (accelerometer, screen state, light, conversation data, and GPS) to identify clusters representing behavior subgroups. Morshed et al [81] developed a computational method to predict mood stability from behavioral features (eg, frequency of conversation, number of location changes, and duration of different physical activities) extracted from accelerometer, microphone, GPS, and Wi-Fi. Recently, de Moura et al [82,83] developed a solution capable of detecting sociability patterns and routine changes in social event streams (ie, conversation events).

A related issue is the predominance of solutions developed for Android OS, for which all apps have a version. This is expected as Android provides an open development platform, different from iOS, with significantly more flexibility to gather the data of interest. The divergent approaches to sensing on iOS and Android yield further issues in terms of standardization and the collection of comparable results across large cohorts, invariably with both Android and iOS users.

Our review further shows that studies use a mix of smartphone-based sensing and wearable device sensing. The latter may be useful where smartphones do not provide quality data (eg, for heart rate, physical activity during sport, or sleep quality), but do pose an issue in terms of interpretability of data given the variety of wearable devices available on the market, each of which use different algorithms. The interpretability of resulting information is further confounded as some of the most popular devices use proprietary algorithms to measure the behaviors of interest or provide aggregate data. Standards would need to consider the commercial pressure for device manufacturers that results in algorithms being proprietary and thus making it difficult to compare information from different devices.

Regarding the year of publication of the studies, most articles (n= 9) have been published in the last 3 years (Figure 7). These data reveal a growing trend in the number of solutions proposed for DPMH.

Research Opportunities

From this review, we are able to identify different research opportunities for DPMH sensing apps, which are open issues for further investigation.

Wearable-Based Solutions

Raw data have been generated mainly in smartphones, so few sensing apps have taken advantage of the potential of wearable devices to produce monitored individual's data ubiquitously. Wearables are capable of providing a lot of useful information about human behavior [79]. For example, wearable devices such as smartwatches and wristbands can collect users' context data even when they are performing intense physical activities such as running and swimming. Therefore, as these devices are smaller, meaning more imperceptible to the user, they can enrich the physiological data collection [85].

Explainable Models With a Focus on Human Behavior

DPMH sensing apps that perform data analysis to design intelligent models have used traditional ML algorithms in different tasks [20]. These models sometimes lack transparency, which is not helpful for mental health professionals because evidence in decision support tools is required to be explainable. Although traditional ML models are very useful for generating valuable information that supports mental health treatment, an explanation of how they generate their outputs is desirable. This is fundamental because professionals need to interpret the patient's behavior to perform assessments and interventions. Therefore, explainable models [86] seems to be the way to apply machine and deep learning techniques more suitable to DPMH.

Real-Time Inference Engines

Most sensing apps perform offline data analysis after collecting raw data (eg, to create ML models, to correlate self-reports with context data). Therefore, few solutions provide inference engines to produce high-level information in real time. These generated situations of interest are useful to have a better insight into the patient's behavior and to allow interventions to adapt to this information in real time. This is crucial in extreme cases such as signs of suicidal ideation, but generally useful where the goal is to implement ecological momentary interventions or just-in-time interventions that rely on just-in-time information on user status. In this sense, both rule-based engines (eg, fuzzy logic [39], complex event processing [82]) and ML-based approaches [20] are promising tools to process context data efficiently and infer high-level information in DPMH.

Extensible Solutions

Sensing apps are not able to be customized for use in other research. Although general-purpose (eg, Sensus [35]) and reusable (eg, Beiwe [15] and SituMan [39]) apps can be applied to other research, none of the solutions identified in this review is extensible. Proposals of framework, middleware, and library are examples of extensible solutions that provide services, reusable code, and are prepared to be modified or consumed by apps. They would be very useful to allow DPMH researchers to extend solution's capabilities to different requirements. Therefore, this could reduce costs and time for research in specific scenarios.

By analyzing the results of the public data set review, we clearly identify the scarcity of data sets (n=8). This low number may be related to the privacy of information collected from study participants. DPMH researchers should possibly be concerned about whether collected data will become public, which could enable to identify participants from them. DPMH data sets may have sensitive personal information about the mental health treatment or monitoring, hence ethical issues arise [87]. Moreover, ethics committees where studies are recorded may restrict the sharing of collected data to the public. This barrier can generate great difficulty for the development of new research, because new ML models and engines for inferring high-level information are not possible to be designed and trained. Differential privacy seems to be a promising tool to break this barrier [88].

Another open issue is the standardization of data sets. Currently, there is no standard for data representation (eg, data type, precision, file format) and collection (eg, frequency, duration, presence of time stamps). As a result, data sets cannot be combined, nor can we easily compare the performance of different approaches or algorithms. Proposals for standardization would be a major contribution to the DPMH field.

It is beneficial for such standardization that there are efforts to design general-purpose sensing apps. We propose that the research community should endeavor to work on such apps collaboratively and make these apps available on a non-for-profit basis. This could not only result in an efficient use of commonly agreed standards, but would also reduce the wasteful effort of developing custom sensing apps. Such initiatives, however, are difficult to start and maintain, as has been shown by brave endeavors such as Beiwe [15], Funf [31], Purple Robot [33], and Sensus [35], which show that keeping such platforms up-to-date is an expensive process that can only be warranted if continued use guarantees continued resources for maintenance and further development.

Notwithstanding the benefits we believe would be derived from such standards, it should be acknowledged that self-reports will likely remain an important modality to improve the quality of automatically measured behaviors, or to measure behaviors or states that cannot be automatically measured. An opportunity that is not widely leveraged is using the automatically measured behaviors to trigger such self-reports. This would allow self-reports to be more appropriate to the user's context, further inform automated measures in case sensor measurements do not provide a clear enough picture, and be less intrusive. Software for such a functionality has been proposed previously [39,89], and we believe such a functionality should be part of standardized tools for capturing DPMH.

Finally, data sets are composed of few study participants. It may be difficult for researchers in attracting participants to the research and, at the same time, making them remain until the end of the study. The low number of participants can potentially compromise the use and validation of some data contained in the data sets, and this directly reflects the use of data sets in other DPMH surveys, where it requires a high number of participants to be validated.

Limitations and Future Work

A first limitation is that data sets and articles published in languages other than English were not included in this review. Second, the search for sensing apps was restricted to 8 digital libraries, although we searched 20 sources with numerous public data sets. Finally, our review is limited by studies reported in the published literature and data sets available to be downloaded.

In addition, we did not focus on security and privacy aspects of DPMH apps in this review. Therefore, our plans include a systematic analysis on the security and privacy features provided by DPMH apps. As this is an extremely sensitive aspect in the development of new functionalities for current and new DPMH mobile systems, a particular characterization with deeper analysis is required. Therefore, we plan to dedicate efforts on this topic for further investigation.

Conclusions

In this article, we described a systematic review that resulted in a deep analysis of 31 sensing apps and 8 public data sets for DPMH. Results showed a growth in DPMH sensing apps in recent years as opposed to a scarcity of public data sets. We answered the research questions, then showing, for example, the most used context data and their respective sources, the different types of high-level information generated by the analysis of the collected data, the features available in data sets, and the mental disorders that researchers have focused. From the results, we were able to identify trends and open issues that hinder the development of research in the DPMH area. As a consequence, by considering the growth in proposals for DPMH sensing apps and the impact of the COVID-19 outbreak on global mental health, we believe that DPMH presents a great perspective for future research not only to overcome open issues discussed in this review, but also to reach the needed maturity for application in clinical settings.

Acknowledgments

This research was funded by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES - Finance code 001); by the Fundação de Amparo à Pesquisa e ao Desenvolvimento Científico e Tecnológico do Maranhão - FAPEMA (via Grants UNIVERSAL-00745/19, BD-01075/20, BEPP-01608/21, BEPP-01783/21, BEPP-01768/21), the state of Maranhão research agency; by the FCT/MCTES through national funds and when applicable co-funded European Union funds under the Project UIDB/50008/2020; and by the Brazilian National Council for Scientific and Technological Development - CNPq, via Grants 313036/2020-9 and 305133/2019-5.

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of 20 data set repositories.
[XLSX File (Microsoft Excel File), 102 KB - jmir_v24i2e28735_app1.xlsx]

Multimedia Appendix 2
Items used in the data extraction process.
[DOCX File , 8 KB - jmir_v24i2e28735_app2.docx]

Multimedia Appendix 3
Reviewed sensing apps.
[DOCX File , 10 KB - jmir_v24i2e28735_app3.docx]

References

1. Ritchie H, Roser M. Mental Health. Our World in Data Internet. URL: https://ourworldindata.org/mental-health [accessed 2021-06-15]
2. World Health Organization. Suicide Data. URL: https://www.who.int/teams/mental-health-and-substance-use/suicide-data [accessed 2021-06-15]
3. World Health Organization. Mental Disorders. URL: https://www.who.int/news-room/fact-sheets/detail/mental-disorders [accessed 2021-06-15]
4. Health at a Glance: Europe 2018: State of Health in the EU Cycle. URL: https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-europe-2018_health_glance_eur-2018-en [accessed 2021-06-15]
5. Mack DL, DaSilva AW, Rogers C, Hedlund E, Murphy EI, Vojdanovski V, et al. Mental Health and Behavior of College Students During the COVID-19 Pandemic: Longitudinal Mobile Smartphone and Ecological Momentary Assessment Study, Part II. J Med Internet Res 2021 Jun 04;23(6):e28892 [FREE Full text] [doi: 10.2196/28892] [Medline: 33900935]
6. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008 Apr 01;4(1):1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
7. Kapur S, Phillips AG, Insel TR. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? Mol Psychiatry 2012 Dec 7;17(12):1174-1179. [doi: 10.1038/mp.2012.105] [Medline: 22869033]
8. Teles A, Rodrigues I, Viana D, Silva F, Coutinho L, Endler M. Mobile Mental Health: A Review of Applications for Depression Assistance. 2019 Presented at: 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS); June 5-7, 2019; Cordoba, Spain p. 708-713. [doi: 10.1109/cbms.2019.00143]
9. Bauer M, Glenn T, Geddes J, Gitlin M, Grof P, Kessing LV, et al. Smartphones in mental health: a critical review of background issues, current status and future concerns. Int J Bipolar Disord 2020 Jan 10;8(1):2-19 [FREE Full text] [doi: 10.1186/s40345-019-0164-x] [Medline: 31919635]
10. Garcia-Ceja E, Riegler M, Nordgreen T, Jakobsen P, Oedegaard KJ, Tørresen J. Mental health monitoring with multimodal sensing and machine learning: A survey. Pervasive and Mobile Computing 2018 Dec;51:1-26. [doi: 10.1016/j.pmcj.2018.09.003]
11. Wind TR, Rijkeboer M, Andersson G, Riper H. The COVID-19 pandemic: The 'black swan' for mental health care and a turning point for e-health. Internet Interv 2020 Apr;20:100317 [FREE Full text] [doi: 10.1016/j.invent.2020.100317] [Medline: 32289019]
12. Torous J, Jän Myrick K, Rauseo-Ricupero N, Firth J. Digital Mental Health and COVID-19: Using Technology Today to Accelerate the Curve on Access and Quality Tomorrow. JMIR Ment Health 2020 Mar 26;7(3):e18848 [FREE Full text] [doi: 10.2196/18848] [Medline: 32213476]
13. Jain SH, Powers BW, Hawkins JB, Brownstein JS. The digital phenotype. Nat Biotechnol 2015 May;33(5):462-463. [doi: 10.1038/nbt.3223] [Medline: 25965751]
14. Liang Y, Zheng X, Zeng DD. A survey on big data-driven digital phenotyping of mental health. Information Fusion 2019 Dec;52:290-307. [doi: 10.1016/j.inffus.2019.04.001]
15. Torous J, Kiang MV, Lorme J, Onnela J. New Tools for New Research in Psychiatry: A Scalable and Customizable Platform to Empower Data Driven Smartphone Research. JMIR Ment Health 2016 May 05;3(2):e16 [FREE Full text] [doi: 10.2196/mental.5165] [Medline: 27150677]
16. Teles A, Barros F, Rodrigues I, Barbosa A, Silva F, Coutinho L. Internet of things applied to mental health: Concepts, applications, and perspectives. In: IoT and ICT for Healthcare Applications. Cham, Switzerland: Springer International Publishing; 2020:33.
17. Tsai C, Lai C, Chao H, Vasilakos AV. Big data analytics: a survey. Journal of Big Data 2015 Oct 1;2(1):1-32. [doi: 10.1186/s40537-015-0030-3]

18. Lan K, Wang D, Fong S, Liu L, Wong KKL, Dey N. A Survey of Data Mining and Deep Learning in Bioinformatics. J Med Syst 2018 Jun 28;42(8):139-120. [doi: 10.1007/s10916-018-1003-9] [Medline: 29956014]
19. Insel TR. Digital Phenotyping: Technology for a New Science of Behavior. JAMA 2017 Oct 03;318(13):1215-1216. [doi: 10.1001/jama.2017.11295] [Medline: 28973224]
20. Benoit J, Onyeaka H, Keshavan M, Torous J. Systematic Review of Digital Phenotyping and Machine Learning in Psychosis Spectrum Illnesses. Harv Rev Psychiatry 2020 Aug 12;28(5):296-304. [doi: 10.1097/hrp.0000000000000268]
21. Cornet VP, Holden RJ. Systematic review of smartphone-based passive sensing for health and wellbeing. J Biomed Inform 2018 Jan;77:120-132 [FREE Full text] [doi: 10.1016/j.jbi.2017.12.008] [Medline: 29248628]
22. De-La-Hoz-Franco E, Ariza-Colpas P, Quero JM, Espinilla M. Sensor-Based Datasets for Human Activity Recognition – A Systematic Review of Literature. IEEE Access 2018;6:59192-59210. [doi: 10.1109/access.2018.2873502]
23. Trifan A, Oliveira M, Oliveira JL. Passive Sensing of Health Outcomes Through Smartphones: Systematic Review of Current Solutions and Possible Limitations. JMIR Mhealth Uhealth 2019 Aug 23;7(8):e12649 [FREE Full text] [doi: 10.2196/12649] [Medline: 31444874]
24. Seppälä J, De Vita I, Jämsä T, Miettunen J, Isohanni M, Rubinstein K, M-RESIST Group, et al. Mobile Phone and Wearable Sensor-Based mHealth Approaches for Psychiatric Disorders and Symptoms: Systematic Review. JMIR Ment Health 2019 Feb 20;6(2):e9819 [FREE Full text] [doi: 10.2196/mental.9819] [Medline: 30785404]
25. Antosik-Wójcińska AZ, Dominiak M, Chojnacka M, Kaczmarek-Majer K, Opara KR, Radziszewska W, Święcicki. Smartphone as a monitoring tool for bipolar disorder: a systematic review including data analysis, machine learning algorithms and predictive modelling. Int J Med Inform 2020 Jun;138:104131. [doi: 10.1016/j.ijmedinf.2020.104131] [Medline: 32305023]
26. Kitchenham B, Charters S. EBSE Technical Report. Guidelines for Performing Systematic Literature Reviews in Software Engineering. 2007. URL: https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.117.471 [accessed 2021-06-15]
27. Perform Systematic Literature Reviews. URL: https://parsif.al/ [accessed 2021-06-15]
28. Wohlin C. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In: Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering - EASE ?. New York, NY: ACM Press; 2014 Presented at: 18th International Conference on Evaluation and Assessment in Software Engineering; May 13-14, 2014; London, UK. [doi: 10.1145/2601248.2601268]
29. Viera A, Garrett J. Understanding interobserver agreement: the kappa statistic. Fam Med 2005;37(5):360-363.
30. Palaghias N, Hoseinitabatabaei SA, Nati M, Gluhak A, Moessner K. A Survey on Mobile Social Signal Processing. ACM Comput. Surv 2016 May 02;48(4):1-52. [doi: 10.1145/2893487]
31. Aharony N, Pan W, Ip C, Khayal I, Pentland A. Social fMRI: Investigating and shaping social mechanisms in the real world. Pervasive and Mobile Computing 2011 Dec;7(6):643-659. [doi: 10.1016/j.pmcj.2011.09.004]
32. Burns MN, Begale M, Duffecy J, Gergle D, Karr CJ, Giangrande E, et al. Harnessing context sensing to develop a mobile intervention for depression. J Med Internet Res 2011 Aug 12;13(3):e55 [FREE Full text] [doi: 10.2196/jmir.1838] [Medline: 21840837]
33. Schueller SM, Begale M, Penedo FJ, Mohr DC. Purple: a modular system for developing and deploying behavioral intervention technologies. J Med Internet Res 2014 Jul 30;16(7):e181 [FREE Full text] [doi: 10.2196/jmir.3376] [Medline: 25079298]
34. Ferreira D, Kostakos V, Dey AK. AWARE: Mobile Context Instrumentation Framework. Front. ICT 2015 Apr 20;2:1-9. [doi: 10.3389/fict.2015.00006]
35. Xiong H, Huang Y, Barnes L, Gerber M. Sensus: A Cross-Platform, General-Purpose System for Mobile Crowdsensing in Human-Subject Studies. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. New York, NY: ACM Press; 2016 Presented at: 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 12-16, 2016; Heidelberg, Germany p. 415-426. [doi: 10.1145/2971648.2971711]
36. Wahle F, Kowatsch T, Fleisch E, Rufer M, Weidt S. Mobile Sensing and Support for People With Depression: A Pilot Trial in the Wild. JMIR Mhealth Uhealth 2016 Sep 21;4(3):e111 [FREE Full text] [doi: 10.2196/mhealth.5960] [Medline: 27655245]
37. Anguera JA, Jordan JT, Castaneda D, Gazzaley A, Areán PA. Conducting a fully mobile and randomised clinical trial for depression: access, engagement and expense. BMJ Innov 2016 Jan 04;2(1):14-21 [FREE Full text] [doi: 10.1136/bmjinnov-2015-000098] [Medline: 27019745]
38. Wang R, Scherer E, Tseng V, Ben-Zeev D, Aung M, Abdullah S. CrossCheck: Toward passive sensing and detection of mental health changes in people with schizophrenia. In: Proceedings of the 20 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp ?. New York, NY: ACM Press; 2016 Presented at: 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 12-16, 2016; New York, NY p. 886-897. [doi: 10.1145/2971648.2971740]
39. Soares Teles A, Rocha A, José da Silva E Silva F, Correia Lopes J, O'Sullivan D, Van de Ven P, et al. Enriching Mental Health Mobile Assessment and Intervention with Situation Awareness. Sensors (Basel) 2017 Jan 10;17(1):127 [FREE Full text] [doi: 10.3390/s17010127] [Medline: 28075417]

40. Servia-Rodríguez S, Rachuri K, Mascolo C, Rentfrow P, Lathia N, Sandstrom G. Mobile Sensing at the Service of Mental Well-being: a Large-scale Longitudinal Study. In: Proceedings of the 26th International Conference on World Wide Web. New York, NY: ACM Press; 2017 Presented at: 26th International Conference on World Wide Web; April 3-7, 2017; Perth, Australia p. 103-112. [doi: 10.1145/3038912.3052618]
41. Wang R, Chen F, Chen Z, Li T, Harari G, Tignor S. StudentLife: Using smartphones to assess mental health and academic performance of college students. In: Mobile Health. Cham, Switzerland: Springer International Publishing; 2017:7.
42. Ben-Zeev D, Scherer EA, Brian RM, Mistler LA, Campbell AT, Wang R. Use of Multimodal Technology to Identify Digital Correlates of Violence Among Inpatients With Serious Mental Illness: A Pilot Study. Psychiatr Serv 2017 Oct 01;68(10):1088-1092 [FREE Full text] [doi: 10.1176/appi.ps.201700077] [Medline: 28669285]
43. Palmius N, Saunders KEA, Carr O, Geddes JR, Goodwin GM, De Vos M. Group-Personalized Regression Models for Predicting Mental Health Scores From Objective Mobile Phone Data Streams: Observational Study. J Med Internet Res 2018 Oct 22;20(10):e10194 [FREE Full text] [doi: 10.2196/10194] [Medline: 30348626]
44. Berrouiguet S, Ramírez D, Barrigón ML, Moreno-Muñoz P, Carmona Camacho R, Baca-García E, et al. Combining Continuous Smartphone Native Sensors Data Capture and Unsupervised Data Mining Techniques for Behavioral Changes Detection: A Case Series of the Evidence-Based Behavior (eB2) Study. JMIR Mhealth Uhealth 2018 Dec 10;6(12):e197 [FREE Full text] [doi: 10.2196/mhealth.9472] [Medline: 30530465]
45. Lind MN, Byrne ML, Wicks G, Smidt AM, Allen NB. The Effortless Assessment of Risk States (EARS) Tool: An Interpersonal Approach to Mobile Sensing. JMIR Ment Health 2018 Aug 28;5(3):e10334 [FREE Full text] [doi: 10.2196/10334] [Medline: 30154072]
46. Chang L, Lu J, Wang J, Chen X, Fang D, Tang Z, et al. SleepGuard. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol 2018 Sep 18;2(3):1-34. [doi: 10.1145/3264908]
47. Elhai JD, Tiamiyu MF, Weeks JW, Levine JC, Picard KJ, Hall BJ. Depression and emotion regulation predict objective smartphone use measured over one week. Personality and Individual Differences 2018 Oct;133:21-28. [doi: 10.1016/j.paid.2017.04.051]
48. Mastoras R, Iakovakis D, Hadjidimitriou S, Charisis V, Kassie S, Alsaadi T, et al. Touchscreen typing pattern analysis for remote detection of the depressive tendency. Sci Rep 2019 Sep 16;9(1):13414 [FREE Full text] [doi: 10.1038/s41598-019-50002-9] [Medline: 31527640]
49. Ranjan Y, Rashid Z, Stewart C, Conde P, Begale M, Verbeeck D, Hyve, RADAR-CNS Consortium. RADAR-Base: Open Source Mobile Health Platform for Collecting, Monitoring, and Analyzing Data Using Sensors, Wearables, and Mobile Devices. JMIR Mhealth Uhealth 2019 Aug 01;7(8):e11734 [FREE Full text] [doi: 10.2196/11734] [Medline: 31373275]
50. Sarda A, Munuswamy S, Sarda S, Subramanian V. Using Passive Smartphone Sensing for Improved Risk Stratification of Patients With Depression and Diabetes: Cross-Sectional Observational Study. JMIR Mhealth Uhealth 2019 Jan 29;7(1):e11041 [FREE Full text] [doi: 10.2196/11041] [Medline: 30694197]
51. Barnett S, Huckvale K, Christensen H, Venkatesh S, Mouzakis K, Vasa R. Intelligent Sensing to Inform and Learn (InSTIL): A Scalable and Governance-Aware Platform for Universal, Smartphone-Based Digital Phenotyping for Research and Clinical Applications. J Med Internet Res 2019 Nov 06;21(11):e16399 [FREE Full text] [doi: 10.2196/16399] [Medline: 31692450]
52. Wisniewski H, Henson P, Torous J. Using a Smartphone App to Identify Clinically Relevant Behavior Trends Symptom Report, Cognition Scores, and Exercise Levels: A Case Series. Front Psychiatry 2019 Sep 23;10:652 [FREE Full text] [doi: 10.3389/fpsyt.2019.00652] [Medline: 31607960]
53. Cao J, Truong AL, Banu S, Shah AA, Sabharwal A, Moukaddam N. Tracking and Predicting Depressive Symptoms of Adolescents Using Smartphone-Based Self-Reports, Parental Evaluations, and Passive Phone Sensor Data: Development and Usability Study. JMIR Ment Health 2020 Jan 24;7(1):e14045 [FREE Full text] [doi: 10.2196/14045] [Medline: 32012072]
54. Narziev N, Goh H, Toshnazarov K, Lee SA, Chung K, Noh Y. STDD: Short-Term Depression Detection with Passive Sensing. Sensors (Basel) 2020 Mar 04;20(5):1396 [FREE Full text] [doi: 10.3390/s20051396] [Medline: 32143358]
55. Dogrucu A, Perucic A, Isaro A, Ball D, Toto E, Rundensteiner EA, et al. Moodable: On feasibility of instantaneous depression assessment using machine learning on voice samples with retrospectively harvested smartphone and social media data. Smart Health 2020 Jul;17:100118. [doi: 10.1016/j.smhl.2020.100118]
56. Betthauser LM, Stearns-Yoder KA, McGarity S, Smith V, Place S, Brenner LA. Mobile App for Mental Health Monitoring and Clinical Outreach in Veterans: Mixed Methods Feasibility and Acceptability Study. J Med Internet Res 2020 Aug 11;22(8):e15506 [FREE Full text] [doi: 10.2196/15506] [Medline: 32779572]
57. Haines-Delmont A, Chahal G, Bruen AJ, Wall A, Khan CT, Sadashiv R, et al. Testing Suicide Risk Prediction Algorithms Using Phone Measurements With Patients in Acute Mental Health Settings: Feasibility Study. JMIR Mhealth Uhealth 2020 Jun 26;8(6):e15901 [FREE Full text] [doi: 10.2196/15901] [Medline: 32442152]
58. Silva E, Aguiar J, Reis LP, Sá JOE, Gonçalves J, Carvalho V. Stress among Portuguese Medical Students: the EuStress Solution. J Med Syst 2020 Jan 02;44(2):45. [doi: 10.1007/s10916-019-1520-1] [Medline: 31897774]
59. Jacobson NC, Chung YJ. Passive Sensing of Prediction of Moment-To-Moment Depressed Mood among Undergraduates with Clinical Levels of Depression Sample Using Smartphones. Sensors (Basel) 2020 Jun 24;20(12):1-16 [FREE Full text] [doi: 10.3390/s20123572] [Medline: 32599801]

60. Masud MT, Mamun MA, Thapa K, Lee D, Griffiths MD, Yang S. Unobtrusive monitoring of behavior and movement patterns to detect clinical depression severity level via smartphone. J Biomed Inform 2020 Mar;103:103371 [FREE Full text] [doi: 10.1016/j.jbi.2019.103371] [Medline: 31935462]
61. Teo J, Davila S, Yang C, Hii A, Pua C, Yap J. Supplementary Data 1 for "Digital phenotyping by consumer wearables identifies sleep-associated markers of cardiovascular disease risk and biological aging". Figshare. 2019. URL: https://tinyurl.com/2p8c44xz [accessed 2021-06-15]
62. Teo JX, Davila S, Yang C, Hii AA, Pua CJ, Yap J, et al. Digital phenotyping by consumer wearables identifies sleep-associated markers of cardiovascular disease risk and biological aging. Commun Biol 2019 Oct 04;2(1):361-310 [FREE Full text] [doi: 10.1038/s42003-019-0605-1] [Medline: 31602410]
63. Garcia-Ceja E, Riegler M, Jakobsen P, Torresen J, Nordgreen T, Oedegaard K. Depresjon Dataset. Zenodo. 2018. URL: https://zenodo.org/record/1219550 [accessed 2021-06-15]
64. Garcia-Ceja E, Riegler M, Jakobsen P, Tørresen J, Nordgreen T, Oedegaard K. Depresjon: A motor activity database of depression episodes in unipolar and bipolar patients. In: Proceedings of the 9th ACM Multimedia Systems Conference. New York, NY: ACM Press; 2018 Presented at: 9th ACM Multimedia Systems Conference; June 12-15, 2018; Amsterdam, the Netherlands p. 472-477. [doi: 10.1145/3204949.3208125]
65. Fryer D. StudentLife data in RData format. Zenodo. 2019. URL: https://zenodo.org/record/3529253 [accessed 2021-06-15]
66. Campbell A. StudentLife Study. URL: https://studentlife.cs.dartmouth.edu/ [accessed 2021-06-15]
67. Wang R, Chen F, Chen Z, Li T, Harari G, Tignor S. StudentLife: Assessing mental health, academic performancebehavioral trends of college students using smartphones. In: Proceedings of the 2014 ACM International Joint Conference on PervasiveUbiquitous Computing. New York, NY. USA: ACM; 2014 Presented at: 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 13-17, 2014; Seattle, Washington p. 3-14. [doi: 10.1145/2632048.2632054]
68. Boonstra T, Nicholas J, Wong Q, Shaw F, Townsend S, Christensen H. Using the Socialise app to collect smartphone sensor data for mental health research: A feasibility study. Zenodo. 2018. URL: https://zenodo.org/record/1238226 [accessed 2021-06-15]
69. Boonstra T, Werner-Seidler A, O?Dea B, Larsen M, Christensen H. Smartphone app to investigate the relationship between social connectivity and mental health. 2017 Presented at: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; July 11-15, 2017; Jeju Island, South Korea p. 287-290. [doi: 10.1109/embc.2017.8036818]
70. Boonstra TW, Nicholas J, Wong QJ, Shaw F, Townsend S, Christensen H. Using Mobile Phone Sensor Technology for Mental Health Research: Integrated Analysis to Identify Hidden Challenges and Potential Solutions. J Med Internet Res 2018 Jul 30;20(7):e10131 [FREE Full text] [doi: 10.2196/10131] [Medline: 30061092]
71. Kiang M, Lorme J, Onnela JP. Public Sample Beiwe Dataset. Zenodo. 2018. URL: https://zenodo.org/record/1188879 [accessed 2021-06-15]
72. Burchert S, Kerber A, Zimmermann J, Knaevelsrud C. 14-day smartphone ambulatory assessment of depression symptoms and mood dynamics in a general population sample: comparison with the PHQ-9 depression screening. Zenodo. 2019. URL: https://zenodo.org/record/3384860 [accessed 2021-06-15]
73. Burchert S, Kerber A, Zimmermann J, Knaevelsrud C. Screening accuracy of a 14-day smartphone ambulatory assessment of depression symptoms and mood dynamics in a general population sample: Comparison with the PHQ-9 depression screening. PLoS One 2021;16(1):e0244955-e0244950 [FREE Full text] [doi: 10.1371/journal.pone.0244955] [Medline: 33406120]
74. Dana Diotte. Sleep Data. Kaggle. URL: https://www.kaggle.com/danagerous/sleep-data [accessed 2021-06-15]
75. Gorgolewski C. Self-tracking. Kaggle. 2018. URL: https://www.kaggle.com/chrisfilo/selftracking [accessed 2021-06-15]
76. Apple Inc.. Sleep Cycle - Sleep Tracker. App Store. URL: https://apps.apple.com/jo/app/%20sleep-cycle-sleep-better/id320606217 [accessed 2021-06-15]
77. Sharp J, Cooper B. Exist. URL: https://exist.io/ [accessed 2021-06-15]
78. Kabadayi S, Pridgen A, Julien C. Virtual sensors: abstracting data from physical sensors. In: Proceedings of the International Symposium on a World of Wireless, Mobile and Multimedia Networks. 2006 Presented at: 2006 International Symposium on a World of Wireless, Mobile and Multimedia Networks; June 26-19, 2006; Buffalo-Niagara Falls, NY p. 587-592. [doi: 10.1109/wowmom.2006.115]
79. Rohani DA, Faurholt-Jepsen M, Kessing LV, Bardram JE. Correlations Between Objective Behavioral Features Collected From Mobile and Wearable Devices and Depressive Mood Symptoms in Patients With Affective Disorders: Systematic Review. JMIR Mhealth Uhealth 2018 Aug 13;6(8):e165 [FREE Full text] [doi: 10.2196/mhealth.9691] [Medline: 30104184]
80. Saeb S, Lattie EG, Schueller SM, Kording KP, Mohr DC. The relationship between mobile phone location sensor data and depressive symptom severity. PeerJ 2016;4:e2537 [FREE Full text] [doi: 10.7717/peerj.2537] [Medline: 28344895]
81. Morshed MB, Saha K, Li R, D'Mello SK, De Choudhury M, Abowd GD, et al. Prediction of Mood Instability with Passive Sensing. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol 2019 Sep 09;3(3):1-21 [FREE Full text] [doi: 10.1145/3351233]

82. de Moura IR, Teles AS, Endler M, Coutinho LR, da Silva E Silva FJ. Recognizing Context-Aware Human Sociability Patterns Using Pervasive Monitoring for Supporting Mental Health Professionals. Sensors (Basel) 2020 Dec 25;21(1):86 [FREE Full text] [doi: 10.3390/s21010086] [Medline: 33375630]
83. Moura I, Teles A, Silva F, Viana D, Coutinho L, Barros F, et al. Mental health ubiquitous monitoring supported by social situation awareness: A systematic review. J Biomed Inform 2020 Jul;107:103454 [FREE Full text] [doi: 10.1016/j.jbi.2020.103454] [Medline: 32562895]
84. Farhan A, Lu J, Bi J, Russell A, Wang B, Bamis A. Multi-view Bi-clustering to Identify Smartphone Sensing Features Indicative of Depression. In: Proceedings of the 2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies. 2016 Presented at: 2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies; June 27-29, 2016; Washington, DC p. 264-273. [doi: 10.1109/chase.2016.27]
85. Seneviratne S, Hu Y, Nguyen T, Lan G, Khalifa S, Thilakarathna K, et al. A Survey of Wearable Devices and Challenges. IEEE Commun. Surv. Tutorials 2017;19(4):2573-2620. [doi: 10.1109/comst.2017.2731979]
86. Adadi A, Berrada M. Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). IEEE Access 2018;6:52138-52160. [doi: 10.1109/access.2018.2870052]
87. Maher NA, Senders JT, Hulsbergen AF, Lamba N, Parker M, Onnela J, et al. Passive data collection and use in healthcare: A systematic review of ethical issues. Int J Med Inform 2019 Sep;129:242-247. [doi: 10.1016/j.ijmedinf.2019.06.015] [Medline: 31445262]
88. Gong M, Xie Y, Pan K, Feng K, Qin A. A Survey on Differentially Private Machine Learning [Review Article]. IEEE Comput. Intell. Mag 2020 May;15(2):49-64. [doi: 10.1109/mci.2020.2976185]
89. van de Ven P, O'Brien H, Henriques R, Klein M, Msetfi R, Nelson J, E-COMPARED Consortium. ULTEMAT: A mobile framework for smart ecological momentary assessments and interventions. Internet Interv 2017 Sep;9:74-81 [FREE Full text] [doi: 10.1016/j.invent.2017.07.001] [Medline: 30135840]

Abbreviations

DPMH: digital phenotyping of mental health
DS-RQs: research questions data sets
EC: exclusion criteria
EMA: ecological momentary assessment
GPS: global positioning system
IC: inclusion criteria
ML: machine learning
PHQ-9: 9-item Patient Health Questionnaire
SA-RQs: research questions for sensing apps

Edited by A Mavragani; submitted 12.03.21; peer-reviewed by C Smith, MA Ahmad, H Mehdizadeh; comments to author 01.05.21; revised version received 20.06.21; accepted 23.12.21; published 17.02.22.

Please cite as:
Mendes JPM, Moura IR, Van de Ven P, Viana D, Silva FJS, Coutinho LR, Teixeira S, Rodrigues JJPC, Teles AS
Sensing Apps and Public Data Sets for Digital Phenotyping of Mental Health: Systematic Review
J Med Internet Res 2022;24(2):e28735
URL: https://www.jmir.org/2022/2/e28735
doi:10.2196/28735
PMID:35175202

Review

Interactive Visualization Applications in Population Health and Health Services Research: Systematic Scoping Review

Jawad Chishtie[1,2,3,4], MSc, MD; Iwona Anna Bielska[5], PhD; Aldo Barrera[6], BS, MSc; Jean-Sebastien Marchand[7], PhD; Muhammad Imran[8], MSc; Syed Farhan Ali Tirmizi[4], MSc, MD; Luke A Turcotte[9], PhD; Sarah Munce[1,2,10,11], PhD; John Shepherd[1], MBA, MSc; Arrani Senthinathan[11], BSc, MSc; Monica Cepoiu-Martin[12], MD, PhD; Michael Irvine[13,14], PhD; Jessica Babineau[15,16], MLIS; Sally Abudiab[1], BSc, MSc; Marko Bjelica[1], MSc; Christopher Collins[17], PhD; B Catharine Craven[1,18], MSc, MD; Sara Guilcher[1,11,19], MScPT, MSc, PhD; Tara Jeji[20], MD; Parisa Naraei[21], PhD; Susan Jaglal[1,2,11,22], PhD

[1]Rehabilitation Sciences Institute, Faculty of Medicine, University of Toronto, Toronto, ON, Canada
[2]Toronto Rehabilitation Institute, University Health Network, Toronto, ON, Canada
[3]Center for Health Informatics, University of Calgary, Calgary, AB, Canada
[4]Alberta Health Services, Edmonton, AB, Canada
[5]McMaster University, Hamilton, ON, Canada
[6]Simon Fraser University, Burnaby, BC, Canada
[7]Universite de Sherbrooke, Quebec, QC, Canada
[8]Allama Iqbal Open University, Islamabad, Pakistan
[9]University of Waterloo, Waterloo, ON, Canada
[10]Department of Occupational Science and Occupational Therapy, University of Toronto, Toronto, ON, Canada
[11]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada
[12]University of Calgary, Calgary, AB, Canada
[13]Department of Mathematics, University of British Columbia, Vancouver, BC, Canada
[14]British Columbia Centre for Disease Control, Vancouver, BC, Canada
[15]Library & Information Services, University Health Network, Toronto, ON, Canada
[16]The Institute for Education Research, University Health Network, Toronto, ON, Canada
[17]Faculty of Science, Ontario Tech University, Oshawa, ON, Canada
[18]KITE, Toronto Rehabilitation Institute, University Health Network, Toronto, ON, Canada
[19]Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, ON, Canada
[20]Ontario Neurotrauma Foundation, Toronto, ON, Canada
[21]Department of Computer Science, Ryerson University, Toronto, ON, Canada
[22]Department of Physical Therapy, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Jawad Chishtie, MSc, MD
Rehabilitation Sciences Institute
Faculty of Medicine
University of Toronto
500 University Ave
Toronto, ON, M5G 1V7
Canada
Phone: 1 6479756965
Fax: 1 416 946 8582
Email: jac161@gmail.com

Abstract

Background: Simple visualizations in health research data, such as scatter plots, heat maps, and bar charts, typically present relationships between 2 variables. Interactive visualization methods allow for multiple related facets such as numerous risk factors to be studied simultaneously, leading to data insights through exploring trends and patterns from complex big health care data.

The technique presents a powerful tool that can be used in combination with statistical analysis for knowledge discovery, hypothesis generation and testing, and decision support.

Objective: The primary objective of this scoping review is to describe and summarize the evidence of interactive visualization applications, methods, and tools being used in population health and health services research (HSR) and their subdomains in the last 15 years, from January 1, 2005, to March 30, 2019. Our secondary objective is to describe the use cases, metrics, frameworks used, settings, target audience, goals, and co-design of applications.

Methods: We adapted standard scoping review guidelines with a peer-reviewed search strategy: 2 independent researchers at each stage of screening and abstraction, with a third independent researcher to arbitrate conflicts and validate findings. A comprehensive abstraction platform was built to capture the data from diverse bodies of literature, primarily from the computer science and health care sectors. After screening 11,310 articles, we present findings from 56 applications from interrelated areas of population health and HSR, as well as their subdomains such as epidemiologic surveillance, health resource planning, access, and use and costs among diverse clinical and demographic populations.

Results: In this companion review to our earlier systematic synthesis of the literature on visual analytics applications, we present findings in 6 major themes of interactive visualization applications developed for 8 major problem categories. We found a wide application of interactive visualization methods, the major ones being epidemiologic surveillance for infectious disease, resource planning, health service monitoring and quality, and studying medication use patterns. The data sources included mostly secondary administrative and electronic medical record data. In addition, at least two-thirds of the applications involved participatory co-design approaches while introducing a distinct category, *embedded research*, within co-design initiatives. These applications were in response to an identified need for data-driven insights into knowledge generation and decision support. We further discuss the opportunities stemming from the use of interactive visualization methods in studying global health; inequities, including social determinants of health; and other related areas. We also allude to the challenges in the uptake of these methods.

Conclusions: Visualization in health has strong historical roots, with an upward trend in the use of these methods in population health and HSR. Such applications are being fast used by academic and health care agencies for knowledge discovery, hypotheses generation, and decision support.

International Registered Report Identifier (IRRID): RR2-10.2196/14019

(J Med Internet Res 2022;24(2):e27534) doi:10.2196/27534

KEYWORDS

interactive visualization; data visualization; secondary health care data; public health informatics; population health; health services research

Introduction

Background

As digital medicine advances, visualization applications in population health increasingly provide ways for researchers and practitioners to explore and communicate findings [1], supporting knowledge discovery from disparate large data sources [2]. Visual analytics (VA) has been defined as the "science of analytics reasoning facilitated by visual interfaces" [3], and it is an interdisciplinary field combining visualization, statistical analysis, and advanced analytics such as machine learning and cognitive sciences [4]. A specific approach within VA is the use of interactive visualization, which Ola and Sedig [2] define as computational tools that store, process, and visually represent data, to facilitate interactive exploration. Interactive visualization increases the potential for big data use in health care by supporting sense making, knowledge discovery, and hypothesis generation [2,5]. Simple visualizations such as scatter plots, heat maps, and bar charts typically present 2 facets of the data, displaying attributes and relationships between 2 variables such as a disease condition and risk factors. Interactive visualization methods allow for presentation of multiple related facets such as risk factors to be studied simultaneously, leading to insights through exploring trends and patterns [2,5].

Population health research involves the study of data related to health outcomes and determinants of population health [6,7], whereas health services research (HSR) studies the health system in relation to access, quality, costs, and patient outcomes [8,9]. Both fields involve the analysis of large secondary data sources such as clinical databases, administrative data sets, and electronic medical records (EMRs) [10-12]. In a prior review, we summarized evidence on VA applications in these interrelated fields of health care [13]; this review complements it by reviewing the evidence on interactive visualization applications in population health and HSR.

Recent systematic reviews have summarized visualization methods in varied areas of health care. Among the most cited reviews, the study by West et al [1] synthesized literature on the use of visualization approaches for exploratory analysis of electronic health records (EHRs). Similarly, another well-cited review by Carrol et al [14] summarized the literature on visualization and analytics tools used in infectious disease epidemiology, particularly in relation to geographic information systems (GIS), molecular epidemiology, and social network analysis methods. Islam et al [15] offered a comprehensive view on data mining and theoretical approaches in health care. Wu et al [16] summarized evidence on visualization and analytic technologies for characterizing evaluation methods in health informatics, an area primarily concerned with clinical care. The

most recent related review by Chung et al [17] focused on visual approaches in mental health care policy and systems. To our knowledge, interactive visualization applications have not been studied as a body of literature separate from data visualization and VA; hence, this review is the first systematic synthesis on the subject.

Rationale for a Companion Review

This companion review is our second synthesis of literature on visualization and analytics tools, techniques, and approaches in population health and HSR. Our first publication focused on VA methods in these areas, where we offered an updated definition of VA in health care as "an approach, method, or application for analytic reasoning, exploration, knowledge discovery, and sense making of complex data, using one or more interactive visual interfaces, employing analytic and visual engines" [13]. As part of VA applications, analytic engines involve advanced machine learning, database querying, and manipulation.

Interactive visualization applications typically engage a front-end visual engine such as Tableau [18], Qlik [19], and PowerBI [20]. Although all VA methods carry a visualization component, which may or may not be interactive, interactive visualization applications typically do not involve or report an analytic component. Hence, this companion review on interactive visualization applications illustrates the state of evidence in population health and HSR, focusing on contemporary methods, approaches, tools, and co-design from real-world use cases. This review will be helpful for health care researchers, practitioners, and decision-makers to understand and adopt visualization-based data analysis.

Objectives

The primary objective of this scoping review is to describe and summarize the evidence on interactive visualization applications, methods, and tools being used in population health and HSR and their subdomains in the last 15 years, from January 1, 2005, to March 30, 2019. Our secondary objective is to describe the use cases, metrics, frameworks used, settings, target audience, goals, and co-design of applications.

Methods

Review Methodology and Protocol

Scoping reviews outline the size and scope of available literature and identify the quality and extent of research evidence [21]. We briefly describe the methodological processes relevant to the second part of the review in this section, whereas further details can be found in the published protocol [22]. We primarily followed the guidance provided by the Joanna Briggs Institute [23], as well as the framework for conducting scoping reviews described by Arksey and O'Malley [24], with improvements suggested by Levac et al [25] and Peters et al [26], while using the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) checklist provided by Tricco et al [27] for reporting. The major steps were as follows: determining the research question, identifying relevant studies, abstracting data, and summarizing and reporting the results. The operational concepts and definitions are presented in Textbox 1.

Textbox 1. Operational concepts and definitions.

Concepts and definitions

- Population health, adapted from Kindig and Stoddart [6] and Kindig [7]
  - "The health outcomes of a group of individuals, including the distribution of such outcomes within the group," includes "health outcomes, patterns of health determinants, and policies and interventions that link these two"
- Health services research, adapted from the Canadian Institutes of Health Research [8] and National Libraries of Medicine filters for health services research [28]
  - Research with the "goal of improving the efficiency and effectiveness of health professionals and the health care system"
  - Access to services
  - Utilization of services
  - Cost of services
- Domains of population health and health services research, adapted from Islam et al [15]
  - Clinical populations include a health condition
  - Epidemiologic includes disease distribution and dynamics
  - Demographic includes population-related characteristics such as age and gender
  - Spatiotemporal includes events over time and space
- Problem categories, based on subject area and the aim or aims of the application
  - Epidemiologic monitoring or surveillance
  - Resources and services monitoring and planning
  - Medication use patterns
  - Visualization methodologies
  - Epidemiologic data exploration
  - Health service monitoring, planning, and quality
  - Patient or care pathways
  - Public or patient communication
- Interactivity, adapted from Ola and Sedig [29] and Pike et al [30]
  - Ability to reflect changes in the visual representation, based on one or more variables available on the analytic interface
  - Tasks such as filtering, determining ranges, and finding anomalies, clusters, and the like by providing menus, dropdowns, and other options on the visualization interface
- Tools
  - Software for developing an application
- Use case
  - Use of the application or method to one or more data sets
- Goal of the application, adapted from Islam et al [15]
  - Whether the application was meant for decision support, knowledge discovery, or both
- Analytic capability, adapted from Islam et al [15]
  - Descriptive or predictive analytics or visual exploration of data
- Functions of the visualization presentations from the Graphic Continuum by Schwabish and Ribecca [31]
  - Spatial
  - Change over time
  - Flow

- Distribution
- Ranking
- Magnitude
- Correlation
- Part to whole

- Co-design, adapted from Ward et al [32]
  - Encompasses the partnership of health workers, patients, and designers who aspire toward change, depending on shared knowledge to achieve "better outcomes or improved efficiency"
  - Whether any participatory approach toward co-design was reported by the authors
  - Embedded research: applications developed in response to an expressed need within a health care organization
- Settings and target audience
  - On the basis of the location of the application developed and the overall objectives of the reported application
  - Categories include academia, government health care units, and industry
- Subject of applications
  - Exploratory word frequency analysis of included articles to yield major subject areas for which applications were developed or any other related finding using a word cloud
- Applications in current use, public availability, innovation, and limitations
  - For ascertaining whether the application could be adapted or replicated in future
  - Public availability to ascertain whether the application was developed for the public

Eligibility Criteria

Eligible articles included peer-reviewed published journal and full conference papers in English related to use cases of interactive visualization in population health and HSR. We included articles on spatiotemporal visualization but excluded articles presenting cartographic methods and tools for GIS because these were outside the scope of the research objectives. Similarly, we did not include articles on human-computer interaction, user design, and articles without a use case. Non–peer-reviewed work such as editorials, conference abstracts, and short articles were excluded. The eligibility criteria are presented in Textboxes 2 and 3.

Textbox 2. Inclusion criteria.

Inclusion criteria

- Peer reviewed journal or full conference papers
- From January 1, 2005, to March 30, 2019
- Population health or health services research related
  - Articles with population level or health services research metrics: incidence, prevalence, events over time, and space, access, cost, utilization, disease or condition distribution, as well as social or multiple determinants of health
- Interactive visualization used for a use case with one or more data sets

Textbox 3. Exclusion criteria.

Exclusion criteria

- Articles not in English
- Editorials, projects, reviews, book chapters, short papers, or reports
- Articles on computer vision and medical imaging
- Studies conducted in clinical settings without a population level or health services component, such as from a single hospital or unit
- Articles on device or sensor data, without a population level or health services research component
- Studies reporting a visual analytics component or analytic engine

Sources of Evidence and Search Strategy

The search strategy, its conceptualization, and steps for operationalization are detailed in the review protocol [22]. The search was externally peer reviewed using the Peer Review of Electronic Search Strategies Guideline [33] and included an extensive list of search terms and their variants to cover all related concepts of population health, HSR, visualization, analytics, and interactivity [22]. The 6 databases searched, their platforms, and results are summarized in Table 1. We further hand searched 10 relevant journals, in addition to internet searches [22]. We used the Covidence (Veritas Health Innovation Ltd) platform for screening citations [34] and EndNote (Clarivate) for reference management [35].

Table 1. Databases and search results (N=14,099).

Database name	Search results, n (%)
MEDLINE (life sciences and biomedicine)	4633 (32.86)
Embase (life sciences and biomedicine)	1880 (13.33)
Web of Science (multidisciplinary)	5396 (38.27)
Ei Compendex (engineering and technology)	1267 (8.99)
IEEE Xplore (engineering and technology)	151 (1.07)
Inspec (engineering and technology)	772 (5.48)

Data Charting and Synthesis of Results

In all, 2 independent reviewers screened articles at each stage of the review, including title and abstract screening, full-text screening, and data abstraction. A third reviewer acted as an arbiter in case of conflicts and for validating the data abstracted for their content and level of detail.

The data abstraction encompassed the major concepts in 6 major themes: (1) study characteristics (country, problem category, settings, and target audience), (2) tools and techniques used, (3) data type and visualization methods, (4) domains of population health and HSR, (5) innovation of the application and its current availability and use, and (6) if the application was co-designed with the target audience.

Results

Selection of Sources and Presentation of Results

We identified 14,099 articles from the 6 databases searched. Given the varied sources of the articles, we adapted the method described by Bramer et al [36] for removing duplicate references using EndNote X9 [35]. Among the 14,099 articles, we considered major citation details and identified, double-checked, and removed 2078 (14.74%) duplicates, comparing the title, identifiers, publication platforms, and abstracts. From the remaining 12,021 articles, another 711 (5.91%) duplicates were removed after importing into Covidence [34]. We excluded 96% (10,819/11,310) of the references during the title and abstract screening and 89% (435/491) of the articles during the full-text screening. We did not find additional articles from reference lists of recent systematic and narrative reviews, hand searches of individual journals, and internet searches. Hence, of the initially identified 14,099 articles, we have summarized 56 (0.39%) for reporting in this review. The reasons for exclusion during the full-text-screening are detailed in the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) diagram (Figure 1), whereas the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews reporting checklist is presented in Multimedia Appendix 1.

We have also summarized our results in a visual format using a publicly accessible Tableau dashboard, a screenshot of which is presented in Figure 2 [37]. The abstracted data and complete workbook are available to support replication, adaptation, and further analysis. Operational concepts for each category and reported theme are detailed in the *Methods* section (Textbox 1).

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart for article selection. VA: visual analytics.

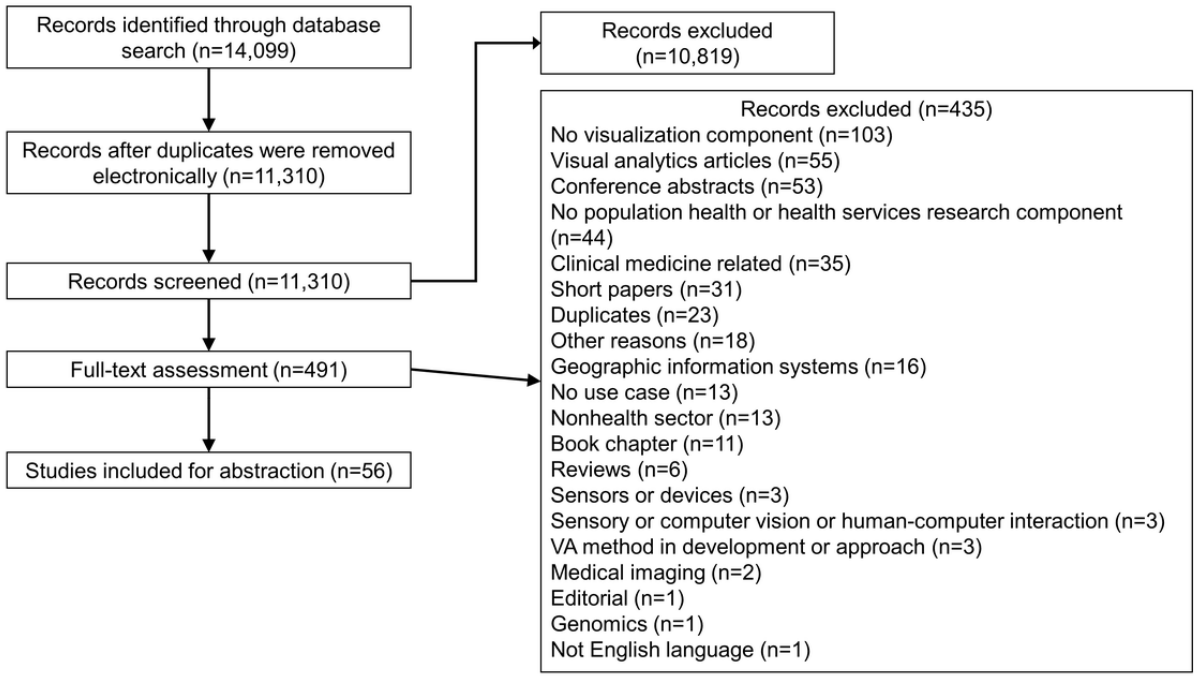
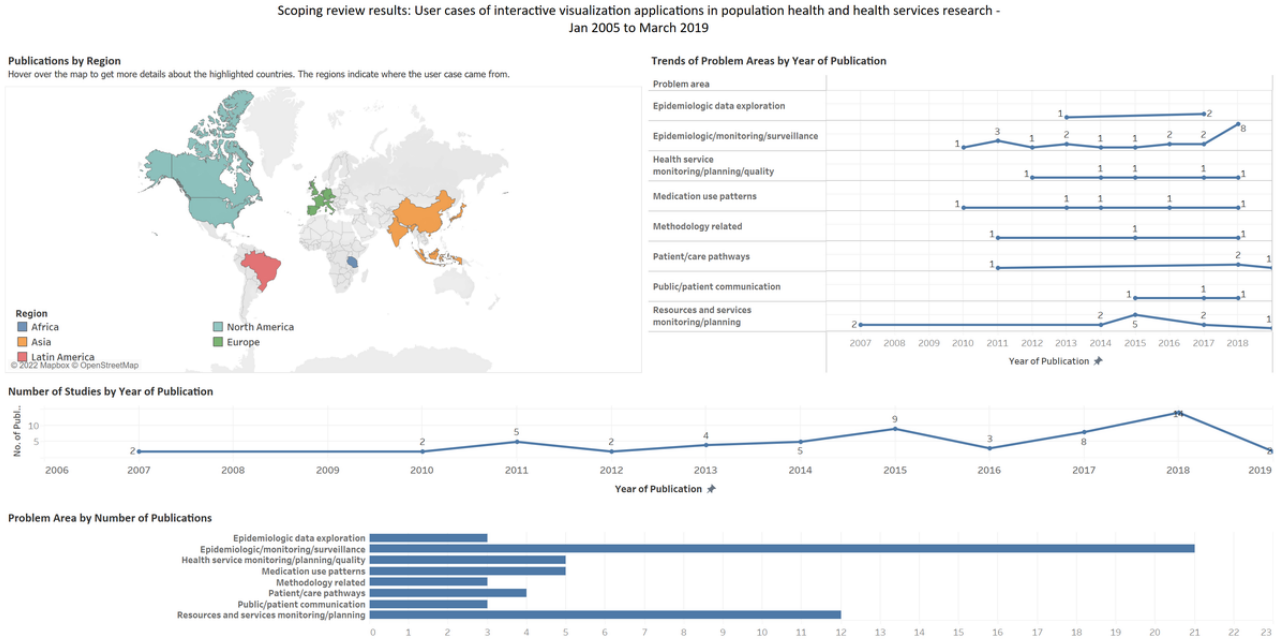


Figure 2. Screenshot of the results presented as a Tableau dashboard.



Study Characteristics, Settings, and Target Audience

The 56 articles summarized were from 21 countries, including the United States (30/56, 54%), the United Kingdom (4/56, 7%), India (2/56, 4%), Indonesia (2/56, 4%), and Canada (2/56, 4%). Of the 56 articles, there was 1 (2%) each from the Netherlands, Spain, Puerto Rico, Czech Republic, Malaysia, France, Portugal, Tanzania, Slovenia, China, Germany, Brazil, Italy, Japan, and Korea, whereas 1 (2%) study included a comparison of health indicators from the United States, the United Kingdom, Costa Rica, Sweden, Croatia, Japan, Hong Kong, and China. Details on countries, settings, and target audiences are presented in Multimedia Appendix 2 [38-93], whereas these are summarized in Tables 2 and 3. Study settings included government ministry or health unit (39/56, 70%), academia (18/56, 32%), and industry (2/56, 4%). There was overlap between the government health unit and academia (1/56, 2%) and between the government health unit and industry (2/56, 4%).

The included studies often had more than one target audience. These were population or public health practitioners (53/56, 95%), clinicians (24/56, 43%), policy makers and decision-makers (21/56, 38%), public and patient groups (12/56, 21%), data scientists (5/56, 9%), and industry (2/56, 4%).

Table 2. Settings of the studies (N=56).

Setting	Values, n (%)	Study
Government; ministry; health department	39 (70)	Alibrahim et al (2014) [38], Barrento and De Castro Neto (2017) [39], Basole et al (2015) [40], BenRamadan et al (2017) [43], BenRamadan et al (2018) [44], Bjarnadottir et al (2016) [46], Brownstein et al (2010) [47], Henley et al (2018) [52], Hosseinpoor et al (2018) [53], Jia et al (2015) [56], Kirtland et al (2014) [58], Ko and Chang (2018) [59], Kubasek et al (2013) [61], Lanzarone et al (2016) [62], Lopez-DeFede et al (2011) [63], Mahler et al (2015) [64], Marshall et al (2017) [65], Mitrpanont et al (2017) [67], Moni et al (2015) [68], Monsen et al (2015) [69], Monsivais et al (2018) [70], Mozumder et al (2018) [71], Pachauri et al (2014)[73], Palmer et al (2019) [74], Pike et al (2017) [76], Podgornik et al (2007) [77], Pur et al (2007) [78], Raghupathi and Raghupathi (2018) [79], Ratwani and Fong (2015) [80], Rodriguez-Fernandez et al (2016) [81], Rowlingson et al (2013) [82], Shen et al (2018) [84], Sims et al (2011) [85], Sopan et al (2012) [86], Toyoda and Niki (2015) [87], Valdiserri and Sullivan (2018) [89], van der Corput et al (2014) [90], Wang and Yao (2018) [92], and Zhang et al (2011) [93]
Academia	18 (32)	Becnel et al (2019) [41], Benítez et al (2017) [42], Bieh-Zimmert et al (2013) [45], Bjarnadottir et al (2016) [46], Cesario et al (2012) [48], Chui et al (2011) [49], Haque et al (2014) [50], Happe and Drezen (2018) [51], Hsu et al (2018) [54], Iyer et al (2017) [55], Kaushal et al (2018) [57], Krause (2015) [60], Martinez et al (2016) [66], Ortiz-Zuazaga et al (2015) [72], Pickle and Carr (2010) [75], Semple et al (2013) [83], Tsoi et al (2018) [88], and Wang et al (2011) [91]
Industry	2 (4)	Ratwani and Fong (2015) [80] and Shen et al (2018) [84]

Table 3. Target audience of the included studies (N=56).

Target audience	Values, n (%)	Study
Population or public health practitioners	53 (95)	Alibrahim et al (2014) [38], Barrento and De Castro Neto (2017) [39], Becnel et al (2019) [41], Benitez et al (2017) [42], BenRamadan et al (2017) [43], BenRamadan et al (2018) [44], Bieh-Zimmert et al (2013) [45], Bjarnadottir et al (2016) [46], Brownstein et al (2010) [47], Cesario et al (2012) [48], Chui et al (2011) [49], Haque et al (2014) [50], Happe and Drezen (2018) [51], Henley et al (2018) [52], Hosseinpoor et al (2018) [53], Hsu et al (2018) [54], Iyer et al (2017) [55], Jia et al (2015) [56], Kaushal et al (2018) [57], Kirtland [58] 2014, Krause (2015) [60], Kubasek et al (2013) [61] , Lopez-DeFede et al (2011) [63], Mahler et al (2015) [64], Marshall et al (2017) [65], Martinez et al (2016) [66], Mitrpanont et al (2017) [67], Moni et al (2015) [68], Monsen et al (2015) [69], Monsivais et al (2018) [70], Mozumder et al (2018) [71], Ortiz-Zuazaga et al (2015) [72], Pachauri et al (2014) [73], Palmer et al (2019) [74], Pickle and Carr (2010) [75], Pike et al (2017) [76], Podgornik et al (2007) [77], Pur et al (2007) [78], Raghupathi and Raghupathi (2018) [79], Ratwani and Fong (2015) [80], Rodriguez-Fernandez et al (2016) [81], Rowlingson et al (2013) [82], Semple et al (2013) [83], Shen et al (2018) [84], Sims et al (2011) [85], Sopan et al (2012) [86], Toyoda and Niki (2015) [87], Tsoi et al (2018) [88], Valdiserri and Sullivan (2018) [89], van der Corput et al (2014) [90], Wang et al (2011) [91], Wang and Yao (2018) [92], and Zhang et al (2011) [93]
Clinicians	24 (43)	Basole et al (2015) [40], Becnel et al (2019) [41], BenRamadan et al (2017) [43], BenRamadan et al (2018) [44], Bjarnadottir et al (2016) [46], Brownstein et al (2010) [47], Haque et al (2014) [50], Happe and Drezen (2018) [51], Henley et al (2018) [52], Jia et al (2015) [56], Kaushal et al (2018) [57], Kirtland et al (2014) [58], Ko and Chang (2018) [59], Lanzarone et al (2016) [62], Marshall et al (2017) [65], Mitrpanont et al (2017) [67], Monsen et al (2015) [69], Mozumder et al (2018) [71], Palmer et al (2019) [74], Pike et al (2017) [76], Ratwani and Fong (2015) [80], Rodriguez-Fernandez et al (2016) [81], Semple et al (2013) [83], and van der Corput et al (2014) [90]
Policy makers and decision-makers	21 (38)	Alibrahim et al (2014) [38], Becnel et al (2019) [41], Hsu et al (2018) [54], Jia et al (2015) [56], Lanzarone et al (2016) [62], Mahler et al (2015) [64], Marshall et al (2017) [65], Moni et al (2015) [68], Monsen et al (2015) [69], Monsivais et al (2018) [70], Pike et al (2017) [76], Podgornik et al (2007) [77], Pur et al (2007) [78], Raghupathi and Raghupathi (2018) [79], Rowlingson et al (2013) [82], Semple et al (2013) [83], Sims et al (2011) [85], Sopan et al (2012) [86], Toyoda and Niki (2015) [87], Valdiserri and Sullivan (2018) [89], Wang (2018) [92], and Zhang et al (2011) [93]
Public and patient groups	12 (21)	Barrento and De Castro Neto (2017) [39], Bieh-Zimmert et al (2013) [45], Brownstein et al (2010) [47], Hosseinpoor et al (2018) [53], Hsu et al (2018) [54], Jia et al (2015) [56], Kubasek et al (2013) [61], Mozumder et al (2018) [71], Ortiz-Zuazaga et al (2015) [72], Semple et al (2013) [83], Tsoi et al (2018) [88], and van der Corput et al (2014) [90]
Data scientists	5 (9)	BenRamadan et al (2017) [43], Pickle and Carr (2010) [75], Tsoi et al (2018) [88], Valdiserri and Sullivan (2018) [89], and Wang et al (2011) [91]
Industry (software, pharmaceutical, and insurance)	2 (4)	Kaushal et al (2018) [57] and Toyoda and Niki (2015) [87]

Health Care Domains, Metrics, and Categories of Problems Addressed by the Applications

Among the domains of health, the categories overlapped, with articles falling under population health (38/56, 68%), HSR (29/56, 52%), and both population health and HSR (11/56, 20%). Among the articles in the population health category, their subdomains included clinical populations with 1 condition of interest (23/56, 41%), demographic population (28/56, 50%), epidemic monitoring and modeling (11/56, 20%), and spatiotemporal (16/56, 29%). For HSR, these included access to services (16/56, 29%), utilization (23/56, 41%), and costs (4/56, 7%).

The visual applications for these health care areas used different metrics in combination with the major categories, including prevalence (23/56, 41%), space and time (20/56, 36%), incidence (19/56, 34%), resources (6/56, 11%), mortality (4/56, 7%), hospitalization (1/56, 2%), events over time (1/56, 2%), and air quality (1/56, 2%).

The problem categories addressed by the applications included epidemiologic monitoring or surveillance (21/56, 38%), resources and services monitoring or planning (12/56, 21%), health service monitoring or planning or quality (5/56, 9%), medication use patterns (5/56, 9%), patient or care pathways (4/56, 7%), visualization methodologies (3/56, 5%), epidemiologic data exploration (3/56, 5%), and public or patient communication (3/56, 5%).

Application's Analytic Capability, Goal, and Frameworks Used

There was overlap in the analytic capability of the tools with applications capable of descriptive analytics (53/56, 95%), predictive analytics (4/56, 7%), and visual exploration of complex data sets (37/56, 66%). Regarding the goal of the visualization application, there was overlap between knowledge discovery (56/56, 100%) and decision support (47/56, 84%). Of the 56 articles, 6 (11%) used a framework in their methods for developing the application. These frameworks are summarized in Table 4. Multimedia Appendix 3 [38-93] lists the analytic capability and goals of each application.

Table 4. Articles mentioning the use of methodological frameworks (N=6).

Author and year	Methodological frameworks used in developing interactive visualization applications
Alibrahim et al (2014) [38]	Display principles for visual monitoring by Few et al [94]
Bieh-Zimmert et al (2013) [45]	Ten guidelines by Kelleher and Wagener [95]
Monsen et al (2015) [69]	Followed the Omaha System [96]
Ratwani et al (2015) [80]	Visualization principles (overview, zoom and filter, and details on demand) based on theories from Shneiderman [97] and Chen [98]
Semple et al (2013) [83]	For developing the web app, the 5-stage user-centered design model described by Kinzie et al [99] was used
Wang et al (2011) [91]	Align, Rank, and Filter Framework used for user interaction by Wang et al [100]

Data Characteristics: Source, Structure, Type, and Use Cases

Data sets used in the visualization applications were single (40/56, 71%) or multiple (16/56, 29%), and they were structured (48/56, 86%) or semistructured (8/56, 14%). The sources of data included administrative (45/56, 80%), spatiotemporal (17/56, 30%), EMR or EHR or medical records (15/56, 27%), registry (10/56, 18%), web or social media (2/56, 4%), and sensor data (1/56, 2%). Multimedia Appendix 4 [38-93] details the data types and sources with the primary tools used to develop the application.

Visualization: Primary Types, Presentation, and Tools

Regarding the functional aspects of the interactive visual presentations, the categories included spatial (31/56, 55%), change over time (9/56, 16%), flow (8/56, 14%), distribution (2/56, 4%), ranking (2/56, 4%), magnitude (2/56, 4%), correlation (1/56, 2%), and part to whole (1/56, 2%).

The primary visual presentations included choropleth map (19/56, 34%), thematic map (10/56, 18%), event timeline (7/56, 13%), network map (4/56, 7%), Sankey diagrams (3/56, 5%), area chart (1/56, 2%), parallel coordinates (1/56, 2%), column bars (1/56, 2%), circular weighted graph (1/56, 2%), line (1/56, 2%), dot strip plot (1/56, 2%), ring map (1/56, 2%), table (1/56, 2%), scatterplot matrix (1/56, 2%), bar (1/56, 2%), histogram (1/56, 2%), arc (1/56, 2%), and heat map (1/56, 2%). The relative distribution of visual presentations and software tools by problem category is provided in Figure 3. For details on the functional types and visual presentations included in each article, please refer to Multimedia Appendix 5 [38-93].

The different visualization software tools used included Tableau (7/56, 13%); D3.JS (5/56, 9%); ArcGIS and Instant Atlas (3/56, 5% each); R/R-Shiny, Open Street Map, Google Maps application programming interface (API), SQL, and Java-based application (2/56, 4% each); and MS Power BI, SigmaJS, RESTful API, CNGI, Lifelines2, AtlasPR, Circos, IBM Watson Analytics, SAS BI, Pajek, Gephi, pChart, Three Table View, Python, and QuantumGIS (1/56, 2% each). Some articles did not mention the visualization tool (13/56, 23%). Figure 4 shows a screenshot from the Tableau results dashboard with the primary visualization tools and heat map of problem category and visual presentation. This interactive dashboard is also available on the Tableau results dashboard [37].

Figure 3. Types of visualizations, primary software tool, and visualization type by problem area (screenshot).

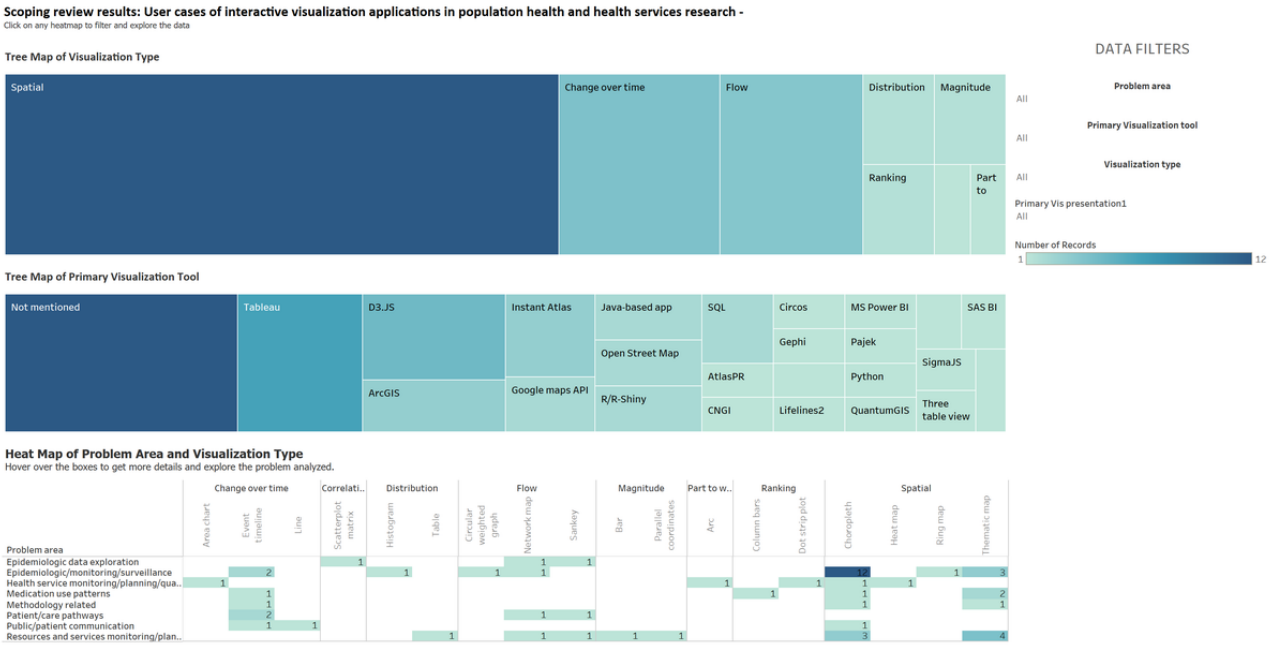


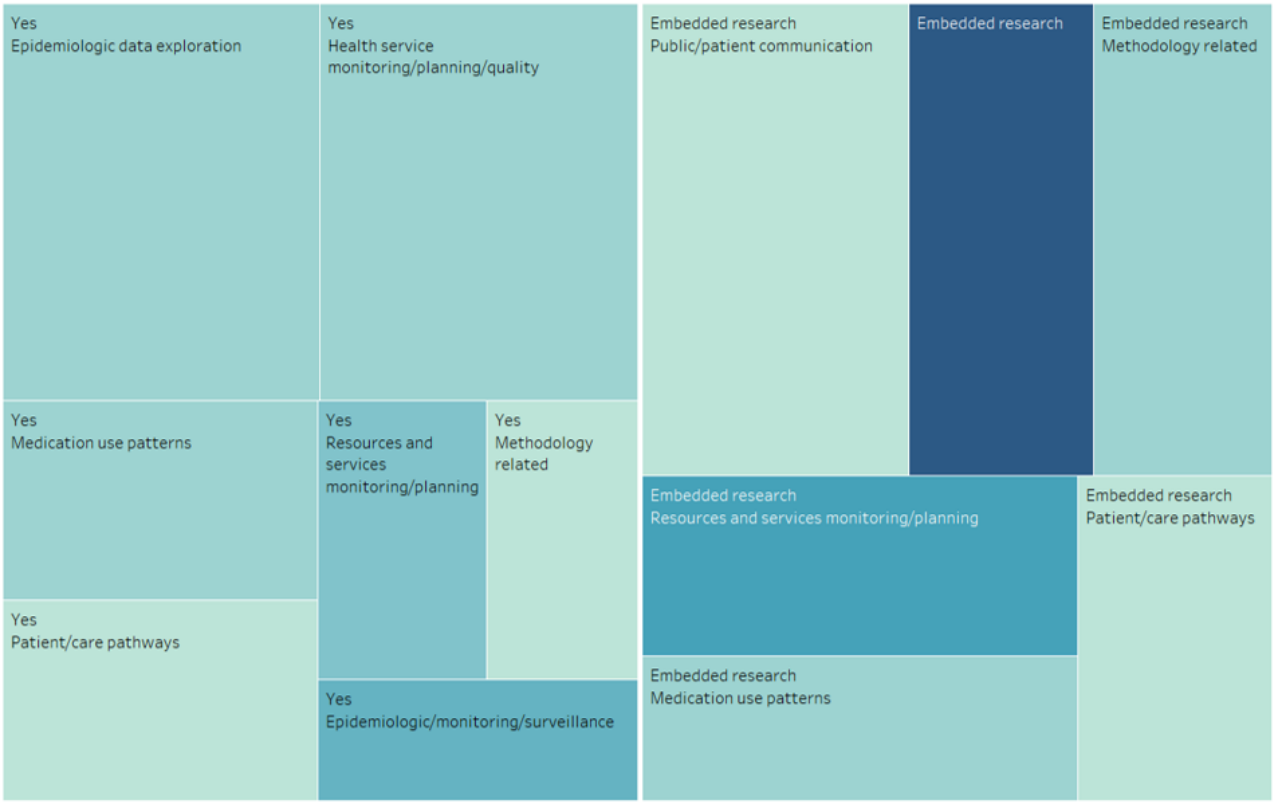
Figure 4. Primary visualization presentations by health care problem area.

Application Co-design

For participatory approaches to application development, some articles (35/56, 67%) mentioned co-design. Among these (20/35, 57%) were applications that were part of embedded research at health care organizations. Other articles (20/56, 37%) did not mention this aspect. Application co-design was found in the problem categories of epidemiologic monitoring or surveillance (13/56, 23%), resource and service monitoring and planning (8/56, 14%), medication use patterns (4/56, 7%), visualization methodology (3/56, 5%), epidemiologic data exploration (2/56, 4%), health service monitoring or planning or quality (2/56,

4%), patient or care pathways (2/56, 4%), and public or patient communication (1/56, 2%). Figure 5 shows a tree map of co-designed applications and embedded research.

Figure 5. Co-designed applications and related health care areas (screenshot).



Applications in Current Use and Public Availability

Most of the applications were mentioned as being currently available and in use (31/56, 55%). Related to public access, a third of the applications were available to the public (18/56, 32%). There were applications using free or open source tools (18/56, 32%) and those using proprietary tools (19/56, 34%), or the tools were not mentioned (18/56, 32%).

Discussion

Significance of the Review

Data visualization in health has a lengthy history going back to the influential work of John Snow and Florence Nightingale in the 19th century. The field of interactive visualization has developed in parallel with computing power and the availability of large, complex health care data sets for diverse audiences such as clinicians, public health researchers, practitioners, and decision-makers [1,41], with considerable progress made in design methodologies [14]. Our review is a novel synthesis and summary of the literature from a vast body of research that had not previously been covered.

In this methodological review, we aim to capture the current state of knowledge and evidence on the topic of *interactive visualization applications* in population health and HSR, distinguishing them from conventional graphical presentations in health care and the related field of VA. We explored areas in population health and HSR to ascertain where these techniques have been used and identified trends and opportunities for the use of these applications.

As population health and health services researchers and practitioners, our perspective and interest in pursuing this research question were based on developing an in-depth understanding of the state of evidence on the use of visualization-based approaches for big health care data analyses. We anticipated that the review would help diverse audiences in population health and HSR learn from practical applications, inform future research endeavors, and help introduce the analytic method to researchers and students. We discuss our findings in this section with these overarching aims, contrasting the findings from previous reviews in other areas of health care using visualization approaches.

Gaps and Opportunities for Application Co-design

Data visualization aims to convey information *at a glance*, although it assumes that the audience has expertise and visual literacy on the subject matter [101]. In their review of visualization-based applications in infectious disease epidemiology, Carrol et al [14] summarize the audience's information needs and learning behavior and point to 3 important barriers to relaying information to target audiences: (1) time constraints, (2) prior knowledge, and (3) cognitive load [14]. Hence, the design process is imperative for an effective application that allows the user to successfully understand the presented data. Various methodologies outlining effective design requirements and experiences from stakeholders to create new products and solutions have been explored [102-104]. For our

scoping review, we opted to use the term *co-design*, which is more commonly used in health care literature, as opposed to *design thinking* and other related terms [13].

In this review, we found that at least two-thirds of the applications involved co-design approaches, involving stakeholders for developing interactive visualization applications. This was in contrast to a smaller proportion of co-designed VA applications (18%), which were mostly prototypes developed by and for data scientists at academic centers [13]. In line with this finding, more than half of the interactive visualization applications were developed in-house within health care organizations. We termed these initiatives *embedded research* as part of co-designed applications to indicate that these were initiated within the organizations in response to an identified data-driven need for knowledge generation and decision support. We could not find such applications in the VA literature [13]. This indicates an important trend because participatory design and development in health has proven to be a key element in better viability and uptake in planning and implementation of services [105,106].

Notably, a third of the articles in this review did not mention a co-design method, which could be due to authors either opting to omit it or because these were covered elsewhere. We recommend that future research indicate whether the application used co-design approaches. It is important to describe the context in adequate detail to appreciate stakeholder needs, experience, and satisfaction. Furthermore, to map and present methods in sufficient detail, we suggest using established frameworks such as the Munzner Nested Process [107] or Design Thinking for Visualization [108] as reporting tools.

Contrasting Interactive Visualization and VA Applications

Through our recent work in studying visualization methods and applications in population health and HSR, we establish that the fields of interactive visualization and VA share communities of practice, methods, and approaches, but they are conceptually separate with important differences. We highlight the major ones here.

We found that interactive visualization applications were initiated by and targeted at researchers and practitioners within government health care organizations tasked with health services delivery, planning, and policy advice. In contrast, most of the VA applications were from and developed for data scientists [13]. In addition, most interactive visualization applications were developed using front-end engines, especially proprietary tools not requiring an advanced knowledge of coding [13]. Most VA methods and applications were prototypes developed using different combinations of tools, with a very small number using proprietary software [13]. Related to theoretical or conceptual frameworks, VA applications offered 13 different frameworks, whereas we could not identify any of these in this review of interactive visualization applications. However, the latter applications mentioned the use of frameworks at different stages of developing the applications. VA applications also expressly mentioned statistical and machine learning techniques as part of the analytic engine, whereas interactive visualization applications mostly used simple descriptive aggregative techniques. In another distinction, most VA applications were prototypes, whereas most interactive visualization applications were developed for knowledge generation and decision support [13].

Both the VA and interactive visualization techniques seem to have originated from North America and Europe [13]. The top 3 countries identified for VA applications were the United States (24/55, 44%), Canada (5/55, 9%), and Germany (3/55, 5%). The top countries for interactive visualization applications were the United States (30/56, 54%), the United Kingdom (4/56, 7%), and Canada and Indonesia (2/56, 4% each). Both our reviews indicated that most of the applications for both methods were descriptive analytics, with an overlap with exploratory analyses of complex data sets (23/55, 42% for VA and 37/56, 66% for interactive visualization), and a small proportion for predictive analytics. The application goals were comparable, with most being knowledge discovery (35/55, 80% for VA and 56/56, 100% for interactive visualization) or decision support (44/55, 80% for VA and 47/56, 84% for interactive visualization), with considerable overlap (29/55, 53% for VA and 47/56, 84% for interactive visualization). The data sets used for both types of applications were single (32/55, 58% for VA and 40/56, 71% for interactive visualization) and structured (40/55, 73% for VA and 48/56, 86% for interactive visualization). There were no unstructured data sets used for interactive visualization applications. Both types of applications used a small number of semistructured data sets (5/55, 9% for VA and 8/56, 14% for interactive visualization).

As population health and HSR are overlapping concepts, many articles in both reviews overlapped with their foci, methods, and the metrics studied. Among the VA articles, almost all (54/55, 98%) had a population health focus, whereas a third (18/55, 33%) were on HSR. There was a smaller overlap among the interactive visualization applications, with approximately two-thirds (38/56, 68%) focusing on population health and approximately half (29/56, 52%) on HSR.

Comparing the subdomains of population health and HSR, the 2 major categories of articles in the VA review focused on spatiotemporal aspects (27/55, 49%) compared with approximately a third (16/56, 29%) for interactive visualization applications. The next largest subdomains in VA included clinical populations focusing on a condition or cluster of conditions (17/55, 31%) or epidemic monitoring and modeling (18/55, 33%). Among the HSR articles for VA, these were mostly for health services' utilization (15/55, 27%), access to care (10/55, 18%), or costs (2/55, 4%). Conversely, in the interactive visualization literature, the most common subdomain for population health was the study of a demographic population (28/56, 50%), followed by a clinical population (23/56, 41%), and epidemic monitoring and modeling (11/56, 20%). There was a similar trend toward the use of both interactive visualization in HSR, with the most common subdomains being health services' utilization (23/56, 41%), followed by access (16/56, 29%) and costs (4/56, 7%).

The categories of problems have important similarities and variations with epidemiologic surveillance for infectious disease being the major category that the applications targeted (38%

for both VA and interactive visualization). The next problem categories for VA applications were medical record pattern identification (20/55, 36%), population health monitoring (9/55, 16%), and health system resource planning (2/55, 4%). For interactive visualization applications, these included resources and services monitoring or planning (12/56, 21%), health service monitoring or quality (5/56, 9%), and medication use patterns (5/56, 9%).

Interactive visualization applications mostly used administrative and EMR or EHR data sources. This can be attributed directly to the availability of data within health care organizations. VA applications were developed using varied data sources, including administrative (19/55, 35%), EMR or EHR (17/55, 31%), spatiotemporal (16/55, 29%), social media (8/55, 15%), and simulation data (6/55, 11%); for interactive visualization applications, the data sources were secondary administrative data (45/56, 80%), social media (2/56, 4%), and sensor data (1/56, 2%).

Comparing tools in current use, about a third (21/55, 38%) of the VA applications were in use at the time of publication, whereas others were either not available or were prototypes. Moreover, a few (7/56, 13%) applications were accessible for public use, while less than a third were developed using free open source tools (13/56, 24%). Among the interactive visualization applications, more than half (31/56, 55%) were mentioned as being in current use, whereas about a third (18/56, 32%) were available to the public, and the same proportion were developed using free or open source tools. There was a greater proportion of use of proprietary tools (19/56, 34%) for interactive visualization applications compared only a 10th of VA applications (5/55, 10%).

The trend for the use of visual presentations was toward the use of different maps in both applications. Choropleth maps were the most frequently used for interactive visualizations (13/56, 24%), followed by thematic maps (10/56, 18%), event timelines (7/56, 13%), and network maps (4/56, 7%). VA applications showed a similar trend with thematic maps (17/55, 31%), timelines (8/55, 15%), and heat or choropleth maps (6/55, 11%). This corresponds to the findings of the review by Chung et al [17] on visualization methods in the area of mental health systems, which indicated that the most common means of presenting data was through maps [17].

Because of the differences in the methods involved in developing the applications, software tools varied greatly. VA tools were a mix of software tools used for the analytic and visual engines, whereas interactive visualization applications reported visual engines alone. However, there were still similarities in the use of tools. Tableau was the most frequently reported tool for interactive visualization applications (7/56, 13%), followed by D3.JS (5/56, 9%); ArcGIS and Instant Atlas (3/56, 5% each); and R/R-Shiny, Open Street Map, Google Maps API, SQL, and Java-based applications (2/56, 4% each). The most common tools found for VA applications were R-based tools (7/55, 13%), followed by D3.JS (4/55, 7%); SQL (4/55, 7%), Java-based tools (3/55, 5%); and Python-based tools, HTML 5, or Google Maps API (2/55, 4% each). Front-end visual engines such as Tableau were used by only 1 VA application in combination with Weka as the analytic engine.

Finally, an issue that we identified in both our reviews was the lack of reporting detail in the articles, which is important for the replicability and adaptation of the methods used in developing applications. We suggest using part of the VA Reporting Checklist that we presented in our previous work on VA, particularly around the details on the *visualization engine* for the standard reporting of interactive visualization applications [13].

Recent Trends of Using Interactive Visualization Methods

Our results showed that thematic mapping, including choropleth maps, was the most common visual presentation across all problem categories of population health and HSR. This was particularly the case for epidemiologic monitoring and surveillance. The recently created COVID-19 dashboards fall into the same category of applications [109]. Mapping also surfaced as a popular method for health resource monitoring, particularly for the planning of health care services [41,56,62,64,67,73,77,78,80,87,110,111].

Among different conditions of interest, a significant number of applications were developed for studying trends in cancer [43,44,52,61,71,88]. Being a worldwide population health issue, the greater use of interactive visualization methods in cancer could be due to the availability of dedicated registries and secondary administrative data [88]. In global health, applications focused on surveillance of communicable diseases [47], outreach campaigns [64], methods to examine health inequalities [53], and effects on health from global climate change [61]. In HSR, 6 applications directly or indirectly highlighted inequities in health, particularly in regard to effective planning and advising policy [43,53,63,70,82,89]. There was 1 article examining social determinants of health in HIV [63].

As two-thirds of the applications were focused on the visual exploration of complex data sets, this indicated a clear trend toward the use of this technique for exploratory analyses. Although most applications were meant for descriptive analytics and visual exploration of complex data sets, of the 56 applications, 4 (7%) were also capable of predictive analytics [39,68,71,93]. The methodological frameworks that were applied to developing the applications pertained to visual monitoring [38], level of detail in visual presentation [80], use of scientific publication visualizations [45], information management [69], user-centered web-based applications [83], and user interaction [91].

Opportunities for Future Applications and Research

Experts highlight the preference of researchers for interactive graphics to facilitate data exploration and abstraction, and they suggest greater, varied learning opportunities from the use of interactive visualization tools [14].

In comparison with standard, traditional statistical analyses, interactive visualization techniques can play an important complementary role through knowledge generation as well as establishing associations and causality. Interactive visualization

methods enable a *data* discourse, leading to in-depth data-driven insights, while having the advantage of improved perception with reducing cognitive load [2,5]. This interplay of direct data manipulation and analysis allows simultaneous study of trends and patterns in the analytic process, while formulating and testing hypotheses [13,112]. Furthermore, these methods are considered apt for studying correlations in high-dimensional data with a large number of time points [112]. This translates into a powerful technique for using big health care data, allowing a deep exploratory dive without an a priori hypothesis to identify data-driven trends and patterns.

In this review, although we observed various applications of interactive visualization, we found limited evidence of its use in global health. Given the massive open access data sets available from agencies such as the World Bank and the World Health Organization (WHO), research can focus on studying a plethora of population health and HSR indicators [113,114]. The WHO's Global Health Observatory provides population health–related data and statistics from 194 member states, particularly on nutrition, virological surveillance, workforce, and health systems, whereas the World Bank's open data repository features macroeconomic and social indicators such as gender and aid effectiveness. The methods can be helpful in ecologic studies, such as those comparing indicators across and within nations.

Related to this is another major opportunity for the use of interactive visualization in studying inequities, especially those rooted in social determinants of health. Although the social determinants of health have become a major focus for investigating structural inequities, we found only 1 article examining related aspects in the HIV sector [63]. Social determinants of health are defined by the WHO as "conditions in which people are born, grow, live, work and age...shaped by the distribution of money, power and resources at global, national and local levels" [115]. This is especially relevant for investigating structural inequities related to issues of access and use based on race, gender, disability, income distribution, and indigenous populations [116,117]. Taking Canada's example, investigating proximal factors for health among indigenous populations is one of the priority areas for improving health care [116]. Furthermore, high-quality Canadian data can be used to investigate inequities to better understand gaps in access and use of services by underserved populations. This can be done through national administrative data sources such as the Canadian National Ambulatory Care Reporting System, Discharge Abstract Database, and Hospital Morbidity Database, which store data for emergency and ambulatory care [118], as well as hospital inpatient discharges and day surgery [119].

Another major opportunity comes from the extension of using multiple data sources for studying patient journeys and care pathways. With the increasing use of EMR and EHR technologies, especially in primary care, there is an opportunity for researching patient populations along the continuum of care. Another such example is from the United Kingdom's Clinical Practice Research Datalink database, which forms the largest collection of anonymized primary care patient records [120]. In Ontario, Canada, the Electronic Medical Record Administrative Data Linked Database offers high-quality linked data for exploring trends and patterns in care and its provision with the advantage of capturing quality of care measures involving prescriptions and investigations [121].

In a recent review on visualization approaches for supporting mental health systems and policy research, Chung et al [17] indicate that there is a gap in studies that influence policy. Although policy was not our main area of focus for this review, the work indicates that there is an opportunity for informing and advising policy based on the use of big data, especially in the important area of mental health services.

Although the potential for the use of interactive visualization tools for bringing together disparate data sources is valued, there are related concerns for data interpretation, quality, accuracy, and handling [1,14]. Meeting the needs of diverse users and interdisciplinary teams as well as promoting the understanding of visual approaches are 2 related and important challenges to be cognizant of [1,14]. Researchers indicate that understanding the value of these techniques among health care organizations and public health agencies is key to realizing the potential of these methods regarding decision support [17].

Implications and Value-add From the Review

Our work is unique in several respects. Complementing our work on VA applications in population health and HSR [13], this review amalgamates the findings from studies on interactive visualization applications, while delineating the literature to construct a holistic picture on the use of visualization approaches in these areas of health care. Interactive visualization is an increasingly popular method, especially for embedded research within health care organizations. Although traditional statistical methods inform causality and associations of various conditions, interactive visualization presents a complementary opportunity for knowledge discovery, hypotheses generation, and decision support using big health care data.

As a novel method, we present findings from both our scoping reviews on VA and interactive visualization in a dynamic, interactive, and visual format using Tableau dashboards [37,122]. In the interest of greater transparency and replicability, we provide the abstraction database with relevant fields for adaptation and further analysis [37].

We highlight opportunities in areas of research that could benefit from visualization-based methods to promote the understanding and uptake of the methods among the communities of research and practice. This work would also prove useful in further developing visualization-related analytic methods.

Limitations

Although there are several important limitations that we are cognizant of in reporting this review, we made extensive efforts to identify relevant literature, delineate the body of literature on interactive visualization applications, incorporated rigor in our methods through all stages, and went through extensive steps toward validation to present our findings.

We cast a wide net in our literature search covering 6 databases, published the study protocol, and had our search strategy externally peer reviewed. However, we may have missed relevant literature residing in subject-specific databases such

as those of digital art, mathematics, geography, and computer science. In addition, our review was limited to peer-reviewed literature from journal articles and full conference papers, and we focused on health care–related databases. We did not include CINAHL and ACM Digital Library because we could not find unique articles, separate from MEDLINE and IEEE Xplore, during the pilot searches.

In addition, in line with the first review on VA, this literature synthesis is limited to articles published between January 1, 2005, and March 30, 2019. We situate and report the review within the same period as the one on VA applications to complement and contrast findings. Many COVID-19–related visualization products that surfaced later are not included in this review for both reasons of feasibility and the subject being extremely specialized and falling under *outbreak analytics*. However, we plan a rapid analysis of COVID-19–related visual products later in the year. Although we describe interactive visualization applications, we allude only briefly to the challenges in the use of these methods because this was beyond the scope of this review.

Conclusions

Visualization in health has strong historical roots. This systematic literature synthesis informs the state of evidence and trends toward the use of interactive visualization methods in the important and interrelated areas of population health and HSR. We note a significant trend in the use of interactive visualization applications being used in health care organizations, which we term *embedded research*. Such applications are being used by academic and health care agencies for knowledge discovery and generation, as well as decision support. Many of these applications have been co-designed with relevant stakeholders. Although we found a wide array of applications in different subdomains of population health and health services, there are multiple opportunities for the use of these methods in investigating global- and national-level indicators and social determinants of health, as well as constructing patient journeys for a holistic picture of the continuum of care.

Acknowledgments

This review is part of the doctoral work under JAC's Canadian Institutes of Health Research Health System Impact Fellowship 2018-19 and 2019-20. JAC, IAB, MCM, MI, AK, JSM, and LT are Health System Impact Fellows cofunded by the Canadian Institutes of Health Research and their host organizations. The Fellows' host organizations are mentioned in the author affiliations. SJ holds the Toronto Rehabilitation Institute Chair at the University of Toronto.

Authors' Contributions

All authors contributed significantly to the conceptualization and reporting of the review. JC, IAB, AB, and JSM mainly wrote and revised the manuscript in consultation with the others. SA, MB and MI contributed to data abstraction. MI and JC validated the abstracted data. JC provided conceptual guidance on the Tableau visual dashboards, whereas AB and SFAT contributed significantly to the co-design, usability testing, and revisions to present the results. All authors reviewed the manuscript and its subsequent revisions.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) reporting checklist.
[PDF File (Adobe PDF File), 129 KB - jmir_v24i2e27534_app1.pdf]

Multimedia Appendix 2
Details of problems analyzed, settings, and target audience.
[DOCX File , 40 KB - jmir_v24i2e27534_app2.docx]

Multimedia Appendix 3
Analytic capability and goals of the application.
[PDF File (Adobe PDF File), 124 KB - jmir_v24i2e27534_app3.pdf]

Multimedia Appendix 4
Data types, sources, and visualization tools.
[XLSX File (Microsoft Excel File), 29 KB - jmir_v24i2e27534_app4.xlsx]

Multimedia Appendix 5

Functional types and details on visual presentations.
[XLSX File (Microsoft Excel File), 17 KB - jmir_v24i2e27534_app5.xlsx]

References

1. West V, Borland D, Hammond W. Innovative information visualization of electronic health record data: a systematic review. J Am Med Inform Assoc 2015 Mar;22(2):330-339 [FREE Full text] [doi: 10.1136/amiajnl-2014-002955] [Medline: 25336597]
2. Ola O, Sedig K. Beyond simple charts: design of visualizations for big health data. Online J Public Health Inform 2016 Dec 28;8(3):e195 [FREE Full text] [doi: 10.5210/ojphi.v8i3.7100] [Medline: 28210416]
3. Thomas J, Cook K. Illuminating the Path: The R&D Agenda for Visual Analytics. National Visualization and Analytics Center, Institute of Electrical and Electronics Engineers: IEEE Computer Society; 2005:10-186.
4. Kerren A, Stasko J, Fekete JD, North C. Information Visualization: Human-Centered Issues and Perspectives. New York City: Springer; 2008.
5. Ola O, Sedig K. Discourse with visual health data: design of human-data interaction. Multimodal Technol Interact 2018 Mar 20;2(1):10. [doi: 10.3390/mti2010010]
6. Kindig D, Stoddart G. What is population health? Am J Public Health 2003 Mar;93(3):380-383. [doi: 10.2105/ajph.93.3.380] [Medline: 12604476]
7. Kindig D. Understanding population health terminology. Milbank Q 2007;85(1):139-161 [FREE Full text] [doi: 10.1111/j.1468-0009.2007.00479.x] [Medline: 17319809]
8. Health services research. Canadian Institutes of Health Research, Government of Canada. 2019. URL: http://www.cihr-irsc.gc.ca/e/48809.html [accessed 2019-03-13]
9. Lohr KN, Steinwachs DM. Health services research: an evolving definition of the field. Health Serv Res 2002 Feb;37(1):15-17. [doi: 10.1111/1475-6773.01020]
10. Martin-Sanchez F, Verspoor K. Big data in medicine is driving big changes. Yearb Med Inform 2014 Aug 15;9:14-20 [FREE Full text] [doi: 10.15265/IY-2014-0020] [Medline: 25123716]
11. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Inf Sci Syst 2014;2:3 [FREE Full text] [doi: 10.1186/2047-2501-2-3] [Medline: 25825667]
12. White SE. A review of big data in health care: challenges and opportunities. Open Access Bioinform 2014 Oct 31:13-18. [doi: 10.2147/oab.s50519]
13. Chishtie JA, Marchand J, Turcotte LA, Bielska IA, Babineau J, Cepoiu-Martin M, et al. Visual analytic tools and techniques in population health and health services research: scoping review. J Med Internet Res 2020 Dec 03;22(12):e17892 [FREE Full text] [doi: 10.2196/17892] [Medline: 33270029]
14. Carroll LN, Au AP, Detwiler LT, Fu T, Painter IS, Abernethy NF. Visualization and analytics tools for infectious disease epidemiology: a systematic review. J Biomed Inform 2014 Oct;51:287-298 [FREE Full text] [doi: 10.1016/j.jbi.2014.04.006] [Medline: 24747356]
15. Islam MS, Hasan MM, Wang X, Germack HD, Noor-E-Alam M. A Systematic Review on Healthcare Analytics: Application and Theoretical Perspective of Data Mining. Healthcare (Basel) 2018 May 23;6(2):1-18 [FREE Full text] [doi: 10.3390/healthcare6020054] [Medline: 29882866]
16. Wu D, Chen A, Manning J, Levy-Fix G, Backonja U, Borland D, et al. Evaluating visual analytics for health informatics applications: a systematic review from the American Medical Informatics Association Visual Analytics Working Group Task Force on Evaluation. J Am Med Inform Assoc 2019 Apr 01;26(4):314-323 [FREE Full text] [doi: 10.1093/jamia/ocy190] [Medline: 30840080]
17. Chung Y, Bagheri N, Salinas-Perez JA, Smurthwaite K, Walsh E, Furst M, et al. Role of visual analytics in supporting mental healthcare systems research and policy: a systematic scoping review. Int J Inform Manag 2020 Feb;50:17-27 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2019.04.012]
18. Tableau. Seattle, USA: Salesforce URL: https://www.tableau.com/ [accessed 2020-01-01]
19. Qlik. Thomas Bravo, LLC, Pennsylvania, USA. 2020. URL: https://www.qlik.com/us/ [accessed 2022-01-14]
20. Power BI. Microsoft. URL: https://powerbi.microsoft.com/en-us/ [accessed 2020-01-01]
21. Grant M, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. Health Info Libr J 2009 Jun;26(2):91-108 [FREE Full text] [doi: 10.1111/j.1471-1842.2009.00848.x] [Medline: 19490148]
22. Chishtie JA, Babineau J, Bielska IA, Cepoiu-Martin M, Irvine M, Koval A, et al. Visual analytic tools and techniques in population health and health services research: protocol for a scoping review. JMIR Res Protoc 2019 Oct 28;8(10):e14019. [doi: 10.2196/14019]
23. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
24. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
25. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69 [FREE Full text] [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]

26. Peters MDJ, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthc 2015 Sep;13(3):141-146. [doi: 10.1097/XEB.0000000000000050] [Medline: 26134548]
27. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
28. Finding and retrieving HSR: tools and databases. National Library of Medicine. URL: https://www.nlm.nih.gov/nichsr/db.html [accessed 2019-03-13]
29. Ola O, Sedig K. The challenge of big data in public health: an opportunity for visual analytics. Online J Public Health Inform 2014;5(3):223 [FREE Full text] [doi: 10.5210/ojphi.v5i3.4933] [Medline: 24678376]
30. Pike WA, Stasko J, Chang R, O'Connell TA. The science of interaction. Inform Visual 2009 Jan 24;8(4):263-274. [doi: 10.1057/ivs.2009.22]
31. Visual vocabulary. Financial Times. 2017. URL: https://github.com/ft-interactive/chart-doctor/blob/master/visual-vocabulary/Visual-vocabulary.pdf [accessed 2021-11-22]
32. Ward ME, De Brún A, Beirne D, Conway C, Cunningham U, English A, et al. Using co-design to develop a collective leadership intervention for healthcare teams to improve safety culture. Int J Environ Res Public Health 2018 Jun 05;15(6):1182-1192 [FREE Full text] [doi: 10.3390/ijerph15061182] [Medline: 29874883]
33. McGowan J, Sampson M, Salzwedel D, Cogo E, Foerster V, Lefebvre C. PRESS Peer review of electronic search strategies: 2015 guideline statement. J Clin Epidemiol 2016 Jul;75:40-46 [FREE Full text] [doi: 10.1016/j.jclinepi.2016.01.021] [Medline: 27005575]
34. Covidence Systematic Review Software. Veritas Health Innovation, Melbourne, Australia. 2019 Jun 29. URL: https://covidence.org/ [accessed 2020-01-01]
35. EndNote X9 ed. Clarivate, Philadelphia, PA. 2013. URL: https://endnote.com/ [accessed 2020-01-01]
36. Bramer WM, Giustini D, De Jonge GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. J Med Library Assoc 2016 Sep 12;104(3):240-243. [doi: 10.5195/jmla.2016.24]
37. Chishtie JA, Barrera A, Tirmizi FA. Interactive Visualization Applications in Population Health and Health Services Research - Jan 2005 to March 2019. Tableau Public. URL: https://public.tableau.com/app/profile/jawad.chishtie/viz/InteractiveVisualizationApplicationsinPopulationHealthandHealthServicesResearch-Jan2005toMarch2019/DataSource [accessed 2022-01-31]
38. Alibrahim A, Wu S, Guerrero E. Performance dashboard for the substance abuse treatment system in Los Angeles. In: Proceedings of the IIE Annual Conference and Expo. 2014 Presented at: IIE Annual Conference and Expo; May 31 - June 3, 2014; Montreal, QC, Canada URL: https://www.proquest.com/openview/b5c296312fd102d4a887a87ca261dfdf/1?pq-origsite=gscholar&cbl=51908
39. Barrento M, De Castro Neto M. European benchmarking in healthcare: Analysis by country with highlight on Portugal. In: Proceedings of the 12th Iberian Conference on Information Systems and Technologies (CISTI). 2017 Presented at: 12th Iberian Conference on Information Systems and Technologies (CISTI); June 21-24, 2017; Lisbon, Portugal. [doi: 10.23919/CISTI.2017.7975899]
40. Basole R, Braunstein M, Kumar V, Park H, Kahng M, Chau DH, et al. Understanding variations in pediatric asthma care processes in the emergency department using visual analytics. J Am Med Inform Assoc 2015 Mar;22(2):318-323. [doi: 10.1093/jamia/ocu016] [Medline: 25656514]
41. Becnel BJ, Pearce PF, Eiswirth E. Geographical mapping of advanced practice registered nurses in Louisiana: leveraging the power of data display. J Am Assoc Nurse Pract 2019 Jan;31(1):8-16. [doi: 10.1097/JXX.0000000000000091] [Medline: 30383550]
42. Benítez JA, Labra JE, Quiroga E, Martín V, García I, Marqués-Sánchez P, et al. A web-based tool for automatic data collection, curation, and visualization of complex healthcare survey studies including social network analysis. Comput Math Methods Med 2017;2017:2579848 [FREE Full text] [doi: 10.1155/2017/2579848] [Medline: 28529537]
43. Ben Ramadan AA, Jackson-Thompson J, Schmaltz CL. Estimated and Interactively Visualized Female Breast Cancer Incidence Rates in Missouri Senate Districts: 2008-2012. Online J Public Health Inform 2017;9(3):e197 [FREE Full text] [doi: 10.5210/ojphi.v9i3.8084] [Medline: 29403576]
44. Ben Ramadan AA, Jackson-Thompson J, Schmaltz CL. Improving Visualization of Female Breast Cancer Survival Estimates: Analysis Using Interactive Mapping Reports. JMIR Public Health Surveill 2018 May 03;4(2):e42 [FREE Full text] [doi: 10.2196/publichealth.8163] [Medline: 29724710]
45. Bieh-Zimmert O, Koschtial C, Felden C. Representing multidimensional cancer registry data. In: Proceedings of the 13th International Conference on Knowledge Management and Knowledge Technologies. 2013 Sep Presented at: 13th International Conference on Knowledge Management and Knowledge Technologies; September 2013; Graz, Austria p. 1-4. [doi: 10.1145/2494188.2494222]
46. Bjarnadóttir MV, Malik S, Onukwugha E, Gooden T, Plaisant C. Understanding adherence and prescription patterns using large-scale claims data. Pharmacoeconomics 2016 Feb;34(2):169-179. [doi: 10.1007/s40273-015-0333-4] [Medline: 26660349]

47. Brownstein JS, Freifeld CC, Chan EH, Keller M, Sonricker AL, Mekaru SR, et al. Information technology and global surveillance of cases of 2009 H1N1 influenza. N Engl J Med 2010 May 06;362(18):1731-1735 [FREE Full text] [doi: 10.1056/NEJMsr1002707] [Medline: 20445186]
48. Cesario M, Jervis M, Luz S, Masoodian M, Rogers B. Time-based geographical mapping of communicable diseases. In: Proceedings of the 16th International Conference on Information Visualisation. 2012 Jul 11 Presented at: 16th International Conference on Information Visualisation; July 11-13, 2012; Montpellier, France p. 118-123. [doi: 10.1109/iv.2012.30]
49. Chui KK, Wenger JB, Cohen SA, Naumova EN. Visual analytics for epidemiologists: understanding the interactions between age, time, and disease with multi-panel graphs. PLoS One 2011 Feb 15;6(2):e14683 [FREE Full text] [doi: 10.1371/journal.pone.0014683] [Medline: 21347221]
50. Haque W, Urquhart B, Berg E, Dhanoa R. Using business intelligence to analyze and share health system infrastructure data in a rural health authority. JMIR Med Inform 2014 Aug 06;2(2):e16 [FREE Full text] [doi: 10.2196/medinform.3590] [Medline: 25599727]
51. Happe A, Drezen E. A visual approach of care pathways from the French nationwide SNDS database - from population to individual records: the ePEPS toolbox. Fundam Clin Pharmacol 2018 Feb;32(1):81-84. [doi: 10.1111/fcp.12324] [Medline: 28921707]
52. Henley SJ, Gallaway S, Singh SD, O'Neil ME, Lunsford NB, Momin B, et al. Lung cancer among women in the United States. J Womens Health (Larchmt) 2018 Nov;27(11):1307-1316 [FREE Full text] [doi: 10.1089/jwh.2018.7397] [Medline: 30312110]
53. Hosseinpoor AR, Schlotheuber A, Nambiar D, Ross Z. Health Equity Assessment Toolkit Plus (HEAT Plus): software for exploring and comparing health inequalities using uploaded datasets. Glob Health Action 2018;11(sup1):1440783 [FREE Full text] [doi: 10.1080/16549716.2018.1440783] [Medline: 29974823]
54. Hsu Y, Cross J, Dille P, Nourbakhsh I, Leiter L, Grode R. Visualization tool for environmental sensing and public health data. In: Proceedings of the ACM Conference Companion Publication on Designing Interactive Systems. 2018 May Presented at: ACM Conference Companion Publication on Designing Interactive Systems; May 2018; Hong Kong, China p. 99-104. [doi: 10.1145/3197391.3205419]
55. Iyer G, DuttaDuwarah S, Sharma A. DataScope: Interactive visual exploratory dashboards for large multidimensional data. In: Proceedings of the 12th Workshop on Visual Analytics in Healthcare. 2017 Presented at: 12th workshop on Visual Analytics in Healthcare; October 01, 2017; Phoenix, Arizona. [doi: 10.1109/vahc.2017.8387496]
56. Jia Y, Wang X, Sedeh R, Wu Q, Cohen N. A web-based application for visualizing the CMS referral network. In: Proceedings of the International Conference on Healthcare Informatics. 2015 Oct 21 Presented at: International Conference on Healthcare Informatics; October 21-23, 2015; Dallas, Texas. [doi: 10.1109/ichi.2015.91]
57. Kaushal K, Kaushik S, Choudhury A, Viswanathan K, Chellappa B, Natarajan S. Patient journey visualizer: a tool for visualizing patient journeys. In: Proceedings of the International Conference on Machine Learning and Data Science (MLDS). 2017 Presented at: International Conference on Machine Learning and Data Science (MLDS); Dec 14-15, 2017; Noida, India. [doi: 10.1109/mlds.2017.19]
58. Kirtland KA, Burrows NR, Geiss LS. Diabetes interactive atlas. Prev Chronic Dis 2014 Feb 06;11:130300 [FREE Full text] [doi: 10.5888/pcd11.130300] [Medline: 24503340]
59. Ko I, Chang H. Interactive data visualization based on conventional statistical findings for antihypertensive prescriptions using National Health Insurance claims data. Int J Med Inform 2018 Aug;116:1-8. [doi: 10.1016/j.ijmedinf.2018.05.003] [Medline: 29887229]
60. Krause DD. Data lakes and data visualization: an innovative approach to address the challenges of access to health care in Mississippi. Online J Public Health Inform 2015;7(3):e225 [FREE Full text] [doi: 10.5210/ojphi.v7i3.6047] [Medline: 26834938]
61. Kubasek M, Hrebicek J, Duek L, Muzik J, Kalina J. SVOD - System for visualizing of oncological data and their semantic enhancement. In: Proceedings of the 10th International Symposium on Environmental Software Systems. 2013 Presented at: 10th International Symposium on Environmental Software Systems (ISESS); October 9-11, 2013; Neusiedl am See, Austria URL: https://link.springer.com/chapter/10.1007/978-3-642-41151-9_56
62. Lanzarone E, Masclet C, Noël F. A multi-user tool for enhancing the daily replanning and control of visits in home care services. Prod Plann Control 2016 Nov 02;28(3):202-219 [FREE Full text] [doi: 10.1080/09537287.2016.1248869]
63. Lòpez-De Fede A, Stewart JE, Hardin JW, Mayfield-Smith K, Sudduth D. Spatial visualization of multivariate datasets: an analysis of STD and HIV/AIDS diagnosis rates and socioeconomic context using ring maps. Public Health Rep 2011;126 Suppl 3:115-126 [FREE Full text] [doi: 10.1177/00333549111260s316] [Medline: 21836744]
64. Mahler H, Searle S, Plotkin M, Kulindwa Y, Greenberg S, Mlanga E, et al. Covering the last kilometer: using GIS to scale-up voluntary medical male circumcision services in Iringa and Njombe Regions, Tanzania. Glob Health Sci Pract 2015 Sep;3(3):503-515 [FREE Full text] [doi: 10.9745/GHSP-D-15-00151] [Medline: 26374807]
65. Marshall BD, Yedinak JL, Goyer J, Green TC, Koziol JA, Alexander-Scott N. Development of a statewide, publicly accessible drug overdose surveillance and information system. Am J Public Health 2017 Nov;107(11):1760-1763. [doi: 10.2105/AJPH.2017.304007] [Medline: 28933938]

66. Martinez R, Ordunez P, Soliz PN, Ballesteros MF. Data visualisation in surveillance for injury prevention and control: conceptual bases and case studies. Inj Prev 2016 Apr;22 Suppl 1:27-33 [FREE Full text] [doi: 10.1136/injuryprev-2015-041812] [Medline: 26728006]
67. Mitrpanont J, Janekitiworapong N, Ongsritrakul S, Varasai S. MedThaiVis: An approach for thai biomedical data visualization. In: Proceedings of the International Conference on Orange Technologies (ICOT). 2017 Presented at: International Conference on Orange Technologies (ICOT); May 23-24, 2017; Johor, Malaysia. [doi: 10.1109/ict-ispc.2017.8075331]
68. Moni MA, Xu H, Liò P. CytoCom: a Cytoscape app to visualize, query and analyse disease comorbidity networks. Bioinformatics 2015 Mar 15;31(6):969-971 [FREE Full text] [doi: 10.1093/bioinformatics/btu731] [Medline: 25380962]
69. Monsen KA, Peterson JJ, Mathiason MA, Kim E, Lee S, Chi C, et al. Data visualization techniques to showcase nursing care quality. Comput Inform Nurs 2015 Oct;33(10):417-426. [doi: 10.1097/CIN.0000000000000190] [Medline: 26468967]
70. Monsivais P, Francis O, Lovelace R, Chang M, Strachan E, Burgoine T. Data visualisation to support obesity policy: case studies of data tools for planning and transport policy in the UK. Int J Obes (Lond) 2018 Dec;42(12):1977-1986 [FREE Full text] [doi: 10.1038/s41366-018-0243-6] [Medline: 30470805]
71. Mozumder SI, Dickman PW, Rutherford MJ, Lambert PC. InterPreT cancer survival: a dynamic web interactive prediction cancer survival tool for health-care professionals and cancer epidemiologists. Cancer Epidemiol 2018 Oct;56:46-52. [doi: 10.1016/j.canep.2018.07.009] [Medline: 30032027]
72. Ortiz-Zuazaga H, Arce-Corretjer R, Solá-Sloan JM, Conde J. SalHUD-A graphical interface to public health data in Puerto Rico. Int J Environ Res Public Health 2015 Dec 22;13(1):1-5 [FREE Full text] [doi: 10.3390/ijerph13010018] [Medline: 26703677]
73. Pachauri A, Trehan S, Jain T, Khetarpal H. Ananya visual analytics system: applications for strengthening healthcare delivery in Bihar, India. In: Proceedings of the IEEE International Conference on Healthcare Informatics. 2014 Sep 15 Presented at: 2014 IEEE International Conference on Healthcare Informatics; September 15-17, 2014; Verona, Italy. [doi: 10.1109/ichi.2014.35]
74. Palmer R, Utley M, Fulop NJ, O'Connor S. Using visualisation methods to analyse referral networks within community health care among patients aged 65 years and over. Health Informatics J 2020 Mar;26(1):354-375 [FREE Full text] [doi: 10.1177/1460458218824717] [Medline: 30736710]
75. Pickle LW, Carr DB. Visualizing health data with micromaps. Spat Spatiotemporal Epidemiol 2010 Jul;1(2-3):143-150. [doi: 10.1016/j.sste.2010.03.007] [Medline: 22749470]
76. Pike I, Smith J, Al-Hajj S, Fuselli P, Macpherson A. The Canadian atlas of child and youth injury: mobilizing injury surveillance data to launch a national knowledge translation tool. Int J Environ Res Public Health 2017 Aug 30;14(9):1-11 [FREE Full text] [doi: 10.3390/ijerph14090982] [Medline: 28867782]
77. Podgornik M, Postema A, English R, Uhde K, Bloom S, Hicksditors. The influenza data summary: a prototype application for visualizing national influenza activity. In: Zeng D, editor. Intelligence and Security Informatics: Biosurveillance. Berlin: Springer; 2007.
78. Pur A, Bohanec M, Lavrac N, Cestnik B. Data Presentation Methods for Monitoring a Public Health-Care System. Berlin, Heidelberg: Springer; 2007.
79. Raghupathi W, Raghupathi V. An empirical study of chronic diseases in the United States: a visual analytics approach. Int J Environ Res Public Health 2018 Mar 01;15(3):1-11 [FREE Full text] [doi: 10.3390/ijerph15030431] [Medline: 29494555]
80. Ratwani RM, Fong A. 'Connecting the dots': leveraging visual analytics to make sense of patient safety event reports. J Am Med Inform Assoc 2015 Mar;22(2):312-317. [doi: 10.1136/amiajnl-2014-002963] [Medline: 25336592]
81. Rodriguez-Fernandez R, Ng N, Susilo D, Prawira J, Bangs MJ, Amiya RM. The double burden of disease among mining workers in Papua, Indonesia: at the crossroads between Old and New health paradigms. BMC Public Health 2016 Sep 08;16:951 [FREE Full text] [doi: 10.1186/s12889-016-3630-8] [Medline: 27609056]
82. Rowlingson B, Lawson E, Taylor B, Diggle PJ. Mapping English GP prescribing data: a tool for monitoring health-service inequalities. BMJ Open 2013 Jan 03;3(1):1-11 [FREE Full text] [doi: 10.1136/bmjopen-2012-001363] [Medline: 23293241]
83. Semple H, Qin H, Sasson C. Development of a web GIS application for visualizing and analyzing community out of hospital cardiac arrest patterns. Online J Public Health Inform 2013;5(2):212 [FREE Full text] [doi: 10.5210/ojphi.v5i2.4587] [Medline: 23923097]
84. Shen AK, Warnock R, Brereton S, McKean S, Wernecke M, Chu S, et al. Influenza vaccination coverage estimates in the fee-for service Medicare beneficiary population 2006 - 2016: Using population-based administrative data to support a geographic based near real-time tool. Hum Vaccin Immunother 2018;14(8):1848-1852 [FREE Full text] [doi: 10.1080/21645515.2018.1462067] [Medline: 29641277]
85. Sims JN, Isokpehi RD, Cooper GA, Bass MP, Brown SD, St John AL, et al. Visual analytics of surveillance data on foodborne vibriosis, United States, 1973-2010. Environ Health Insights 2011;5:71-85 [FREE Full text] [doi: 10.4137/EHI.S7806] [Medline: 22174586]
86. Sopan A, Noh AS, Karol S, Rosenfeld P, Lee G, Shneiderman B. Community health map: a geospatial and multivariate data visualization tool for public health datasets. Govern Inform Q 2012 Apr;29(2):223-234. [doi: 10.1016/j.giq.2011.10.002]
87. Toyoda S, Niki N. Visualization-based medical expenditure analysis support system. Annu Int Conf IEEE Eng Med Biol Soc 2015 Aug;2015:1600-1603. [doi: 10.1109/EMBC.2015.7318680] [Medline: 26736580]

88. Tsoi K, Chan F, Hirai H, Keung G, Kuo Y, Tai S. Data visualization with IBM Watson Analytics for global cancer trends comparison from World Health Organization. Int J Healthc Inform Syst Inform 2018;13(1):45-54 [FREE Full text] [doi: 10.4018/IJHISI.2018010104]
89. Valdiserri RO, Sullivan PS. Data visualization promotes sound public health practice: the AIDSvu example. AIDS Educ Prev 2018 Feb;30(1):26-34. [doi: 10.1521/aeap.2018.30.1.26] [Medline: 29481299]
90. van der Corput P, Arends J, van Wijk JJ. Visualization of medicine prescription behavior. Comput Graph Forum 2014 Jul 12;33(3):161-170. [doi: 10.1111/cgf.12372]
91. Wang TD, Wongsuphasawat K, Plaisant C, Shneiderman B. Extracting insights from electronic health records: case studies, a visual analytics process model, and design recommendations. J Med Syst 2011 Oct;35(5):1135-1152. [doi: 10.1007/s10916-011-9718-x] [Medline: 21541691]
92. Wang K, Yao L. Terminology enabled spatio-temporal analysis visualization for preterm birth data in the US. In: Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies. 2018 Presented at: 11th International Joint Conference on Biomedical Engineering Systems and Technologies; January 2018; Funchal, Madeira, Portugal. [doi: 10.5220/0006647505100518]
93. Zhang A, Qi Q, Jiang L. Design and implementation of a web-based disease control and emergency response system on CNGI for Mekong subregion: IEEE Computer Society. In: Proceedings IEEE International Conference on Spatial Data Mining and Geographical Knowledge Services. 2011 Presented at: IEEE International Conference on Spatial Data Mining and Geographical Knowledge Services; June 29 - July 1, 2011; Fuzhou, China Fuzhou, China. [doi: 10.1109/icsdm.2011.5969044]
94. Few S. Information Dashboard Design: The Effective Visual Communication of Data. California: O'Reilly Media, Inc; 2006:1-211.
95. Kelleher C, Wagener T. Ten guidelines for effective data visualization in scientific publications. Environ Model Softw 2011 Jun;26(6):822-827. [doi: 10.1016/j.envsoft.2010.12.006]
96. Martin KS. The Omaha System: A Key to Practice, Documentation, and Information Management. Philadelphia: Saunders; 2005:1-484.
97. Shneiderman B. The eyes have it: a task by data type taxonomy for information visualizations. In: Proceedings IEEE Symposium on Visual Languages.: IEEE; 1996 Presented at: IEEE Symposium on Visual Languages; Sept. 3-6, 1996; Boulder, CO, USA. [doi: 10.1109/vl.1996.545307]
98. Chen C. Information visualization. WIREs Comp Stat 2010 Apr 29;2(4):387-403. [doi: 10.1002/wics.89]
99. Kinzie M, Cohn W, Julian M, Knaus W. A user-centered model for web site design: needs assessment, user interface design, and rapid prototyping. J Am Med Inform Assoc 2002;9(4):320-330 [FREE Full text] [doi: 10.1197/jamia.m0822] [Medline: 12087113]
100. Wang T, Plaisant C, Quinn A, Stanchak R, Murphy S, Shneiderman B. Aligning temporal data by sentinel events: discovering patterns in electronic health records. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2008 Presented at: SIGCHI Conference on Human Factors in Computing Systems; April 5 - 10, 2008; Florence Italy. [doi: 10.1145/1357054.1357129]
101. Cairo A. How Charts Lie: Getting Smarter about Visual Information. New York City: WW Norton & Company; 2019.
102. Spence B. Human Aspects of Visualization. Berlin, Heidelberg: Springer; 2011.
103. Pink S, Morgan J. Short-term ethnography: intense routes to knowing. Symbol Interact 2013 Jun 13;36(3):351-361. [doi: 10.1002/symb.66]
104. Millen D. Rapid ethnography: time deepening strategies for HCI field research. In: Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques. 2000 Presented at: 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques; August 17 - 19, 2000; New York USA p. 280-286. [doi: 10.1145/347642.347763]
105. McEvoy R, Tierney E, MacFarlane A. 'Participation is integral': understanding the levers and barriers to the implementation of community participation in primary healthcare: a qualitative study using normalisation process theory. BMC Health Serv Res 2019 Jul 23;19(1):515 [FREE Full text] [doi: 10.1186/s12913-019-4331-7] [Medline: 31337391]
106. Steyn PS, Cordero JP, Gichangi P, Smit JA, Nkole T, Kiarie J, et al. Participatory approaches involving community and healthcare providers in family planning/contraceptive information and service provision: a scoping review. Reprod Health 2016 Jul 22;13(1):88 [FREE Full text] [doi: 10.1186/s12978-016-0198-9] [Medline: 27449128]
107. Sedlmair M, Meyer M, Munzner T. Design study methodology: reflections from the trenches and the stacks. IEEE Trans Visual Comput Graphics 2012 Dec;18(12):2431-2440. [doi: 10.1109/tvcg.2012.213]
108. Machuca B, Ahkin A. School of Interactive Arts and Technology. 2020. URL: https://summit.sfu.ca/item/20497 [accessed 2020-06-30]
109. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Diseas 2020 May;20(5):533-534. [doi: 10.1016/s1473-3099(20)30120-1]
110. Haque W, Urquhart B, Berg E. Visualization of services availability: Building healthy communities. New York City: Springer, Cham; 2014.

111. Basole R, Park H, Gupta M, Braunstein M, Chau D, Thompson M. A visual analytics approach to understanding care process variation and conformance. Proc 2015 Workshop Vis Anal Healthc (2015) 2015 Oct;2015:1-20 [FREE Full text] [doi: 10.1145/2836034.2836040] [Medline: 29177250]
112. Weber G, Hauser H. Interactive Visual Exploration and Analysis. New York: Springer; 2014:161-173.
113. The Global Health Observatory. World Health Organization. 2020. URL: https://www.who.int/data/gho [accessed 2020-10-01]
114. World Bank open data, free and open access to global development data. World Bank. 2020. URL: https://data.worldbank.org/ [accessed 2020-12-31]
115. Solar O, Irwin A. A conceptual framework for action on the social determinants of health. WHO Document Production Services. 2010. URL: https://www.who.int/sdhconference/resources/ConceptualframeworkforactiononSDH_eng.pdf [accessed 2020-10-10]
116. Reading C, Wien F. Health inequalities and social determinants of Aboriginal peoples' health. National Collaborating Centre for Aboriginal Health. 2009. URL: https://www.ccnsa-nccah.ca/docs/determinants/RPT-HealthInequalities-Reading-Wien-EN.pdf [accessed 2020-10-10]
117. Kelly M, Morgan A, Bonnefoy J, Butt J, Bergman V, Mackenbach J. The social determinants of health: Developing an evidence base for political action. National Institutes of Health and Clinical Excellence. 2007. URL: https://www.who.int/social_determinants/resources/mekn_report_10oct07.pdf [accessed 2020-10-10]
118. Emergency and ambulatory care. Canadian Institute for Health Information. URL: https://www.cihi.ca/en/emergency-and-ambulatory-care [accessed 2020-10-10]
119. CIHI. Discharge Abstract Database (DAD) metadata. Canadian Institute for Health Information. URL: https://www.cihi.ca/en/discharge-abstract-database-metadata [accessed 2020-10-10]
120. Williams T, van Staa T, Puri S, Eaton S. Recent advances in the utility and use of the General Practice Research Database as an example of a UK Primary Care Data resource. Ther Adv Drug Saf 2012 Apr;3(2):89-99 [FREE Full text] [doi: 10.1177/2042098611435911] [Medline: 25083228]
121. Tu K, Mitiku TF, Ivers NM, Guo H, Lu H, Jaakkimainen L, et al. Evaluation of Electronic Medical Record Administrative data Linked Database (EMRALD). Am J Manag Care 2014;20(1):15-21. [Medline: 24669409]
122. Jawad CA. Interactive results for the scoping review on visual analytic tools and techniques in population health and health services research. University of Toronto. 2020. URL: https://public.tableau.com/profile/jawad.chishtie#!/ [accessed 2020-10-10]

Abbreviations

API: application programming interface
EHR: electronic health record
EMR: electronic medical record
GIS: geographic information system
HSR: health services research
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
VA: visual analytics
WHO: World Health Organization

Edited by R Kukafka, G Eysenbach; submitted 27.01.21; peer-reviewed by I Pike, A Rezaei Aghdam; comments to author 29.03.21; revised version received 27.04.21; accepted 08.10.21; published 18.02.22.

Please cite as:
Chishtie J, Bielska IA, Barrera A, Marchand JS, Imran M, Tirmizi SFA, Turcotte LA, Munce S, Shepherd J, Senthinathan A, Cepoiu-Martin M, Irvine M, Babineau J, Abudiab S, Bjelica M, Collins C, Craven BC, Guilcher S, Jeji T, Naraei P, Jaglal S
Interactive Visualization Applications in Population Health and Health Services Research: Systematic Scoping Review
J Med Internet Res 2022;24(2):e27534
URL: https://www.jmir.org/2022/2/e27534
doi:10.2196/27534
PMID:35179499

Review

Using Social Media to Engage Knowledge Users in Health Research Priority Setting: Scoping Review

Surabhi Sivaratnam[1,2], BHSc; Kyobin Hwang[2,3]; Alyssandra Chee-A-Tow[4], MPH; Lily Ren[5], MI; Geoffrey Fang[6]; Lindsay Jibb[2,7], RN, PhD

[1]Michael G Degroote School of Medicine, McMaster University, Hamilton, ON, Canada
[2]Child Health Evaluative Sciences, Hospital for Sick Children, Toronto, ON, Canada
[3]Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada
[4]Faculty of Dentistry, University of Toronto, Toronto, ON, Canada
[5]Lane Medical Library, Stanford University, Stanford, CA, United States
[6]Faculty of Applied Science and Engineering, University of Toronto, Toronto, ON, Canada
[7]Lawrence S Bloomberg Faculty of Nursing, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Lindsay Jibb, RN, PhD
Child Health Evaluative Sciences
Hospital for Sick Children
686 Bay Street
Toronto, ON
Canada
Phone: 1 416 813 7654 ext 309160
Email: lindsay.jibb@sickkids.ca

Abstract

Background: The need to include individuals with lived experience (ie, patients, family members, caregivers, researchers, and clinicians) in health research priority setting is becoming increasingly recognized. Social media–based methods represent a means to elicit and prioritize the research interests of such individuals, but there remains sparse methodological guidance on how best to conduct these social media efforts and assess their effectiveness.

Objective: This review aims to identify social media strategies that enhance participation in priority-setting research, collate metrics assessing the effectiveness of social media campaigns, and summarize the benefits and limitations of social media–based research approaches, as well as recommendations for prospective campaigns.

Methods: We searched PubMed, Embase, Cochrane Library, Scopus, and Web of Science from database inception until September 2021. Two reviewers independently screened all titles and abstracts, as well as full texts for studies that implemented and evaluated social media strategies aimed at engaging knowledge users in research priority setting. We subsequently conducted a thematic analysis to aggregate study data by related codes and themes.

Results: A total of 23 papers reporting on 22 unique studies were included. These studies used Facebook, Twitter, Reddit, websites, video-calling platforms, emails, blogs, e-newsletters, and web-based forums to engage with health research stakeholders. Priority-setting engagement strategies included paid platform–based advertisements, email-embedded survey links, and question-and-answer forums. Dissemination techniques for priority-setting surveys included snowball sampling and the circulation of participation opportunities via internal members' and external organizations' social media platforms. Social media campaign effectiveness was directly assessed as number of clicks and impressions on posts, frequency of viewed posts, volume of comments and replies, number of times individuals searched for a campaign page, and number of times a hashtag was used. Campaign effectiveness was indirectly assessed as numbers of priority-setting survey responses and visits to external survey administration sites. Recommendations to enhance engagement included the use of social media group moderators, opportunities for peer-to-peer interaction, and the establishment of a consistent tone and brand.

Conclusions: Social media may increase the speed and reach of priority-setting participation opportunities leading to the development of research agendas informed by patients, family caregivers, clinicians, and researchers. Perceived limitations of the approach include underrepresentation of certain demographic groups and addressing such limitations will enhance the inclusion of diverse research priority opinions in future research agendas.

(J Med Internet Res 2022;24(2):e29821) doi:10.2196/29821

KEYWORDS

social media; research priority-setting; knowledge user; scoping review

Introduction

Background

The need to meaningfully engage individuals with lived experience (ie, patients, family members, caregivers, clinicians, researchers, and other advocates; henceforth referred to as knowledge users) in the conduct of health research—defined as research that includes clinical and basic medical sciences, such as care-based research, systems research, and preventative research—is being increasingly recognized by the scientific community. In particular, it is recognized that these individuals should be included at the onset of the research process, with the aim of developing research that meets the needs of individuals with lived experiences [1]. In fact, the lack of involvement of these individuals in such research priority setting has been identified as a key contributor to difficulties in effectively translating research findings into clinical practice and policy [2].

In parallel, the use of social media—defined as any web-based platform or mobile app through which users can engage with others—is gaining considerable traction within the research community, as researchers increasingly access Facebook, Twitter, and YouTube to support participant recruitment and other research activities [3]. The benefits of research-related social media use include enhanced connectivity between researchers and participants and the potential for rapid diffusion of scientific knowledge to target audiences [4]. The nature of web-based survey methods may also enhance anonymity for participants within the research process, potentially promoting the collection of more valid data [5]. Particularly, data collected via the web may be less vulnerable to contextual biases that can arise in focus group settings or when researchers administer surveys in-person [5].

In light of such potential benefits, a growing body of literature describing the use of social media to elicit and prioritize research uncertainties from knowledge users is emerging [6]. However, there remains sparse methodological guidance on how best to conduct social media efforts and their corresponding effectiveness in developing knowledge user–built research agendas [7].

Objective and Research Questions

Through this knowledge user–driven scoping review, we aim to identify studies that implemented and evaluated social media campaigns that promote participation in setting priorities for health research to address three overarching research questions:

1. What social media–based strategies have been used to enhance knowledge user participation in health research priority setting?
2. What metrics (direct and indirect) have been used to assess the effectiveness of these social media campaigns in securing knowledge user participation?
3. From the perspectives of those conducting social media–based research priority-setting campaigns, what are the benefits and limitations of the method, as well as recommendations for future campaigns?

Methods

Overview

An internal protocol was developed for this review. Our reporting process was conducted in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for scoping reviews) guidelines [8].

Search Strategy and Selection of Studies

A comprehensive search strategy was developed in consultation with a tertiary hospital librarian (LR). We conducted tailored searches in PubMed, Embase, Cochrane Library, Scopus, and Web of Science. We searched all databases from their inception to September 14, 2021. Multimedia Appendix 1 shows the search strategy. Intradatabase and interdatabase duplicates were removed electronically. Using Covidence (Veritas Health Innovation), titles and abstracts were screened independently by 2 trained authors (KH and SS) according to our eligibility criteria. In cases of conflicting opinions on eligibility, studies were moved to full-text screening. Full-text articles were then screened independently by 2 authors (KH and SS). Any eligibility disagreements were resolved by consensus through discussion by at least three authors (AC, KH, SS, and LJ). The reference lists of relevant studies were also scanned to find other applicable papers.

Selection Criteria

We included studies (1) that discussed strategies to promote social media–based health research priority setting among key stakeholders and knowledge users and (2) measured the effectiveness of such strategies directly or indirectly. There were no restrictions on the language, country, and year of publication, nor the research content focus, as priority-setting research is cross-disciplinary. Although no explicit restrictions were placed on the language, the included studies were dominated by English language–based social media campaigns. We defined social media as any web-based platform or mobile app through which users can interact and engage with others. We defined knowledge users as patients (or potential patients), caregivers, clinicians, and other advocates (eg, health researchers). We excluded (1) studies where the purpose of the social media campaign did not include knowledge user engagement (eg, social campaigns used to disseminate smoking cessation information to knowledge users) [9]; (2) studies where the research prioritization campaign did not involve social media (dissemination techniques solely involved telephone calls, flyer distribution, etc); and (3) abstracts, dissertations, protocols, systematic reviews, scoping reviews, or case studies.

Data Extraction and Management

A standard electronic data collection form was created and piloted with our group, after which data extraction occurred independently (KH and SS). Discrepancies between the collected data were resolved through discussion with 3 authors (LJ, SS, and KH).

Data Analyses

We used descriptive statistics to summarize quantitative study data and an inductive thematic analysis to synthesize qualitative data [10]. Our data collection form was uploaded to NVivo (version 12.6.0; QSR International) for analysis and was read through multiple times by 2 authors (KH and SS) who had previous experience with thematic analyses. One author (SS) then coded qualitative text within the table on a segment-by-segment basis. At frequent meetings, a second author (KH) reviewed the coding decisions using a constant comparative approach adapted from Thorne [11]. As a group, we (KH, SS, and LJ) then collapsed these codes into subthemes and themes based on the between-code relationships and in accordance with our research questions.

Results

Overview

Figure 1 outlines our study identification process. Overall, 23 papers reporting on 22 unique studies were included in this review. The number of published studies increased steadily over time until 2020, which was the last complete publication year (Figure 2).

Included studies were conducted in 46 countries, most commonly in the United States (11/23, 48%), the United Kingdom (7/23, 30%), and Canada (5/23, 22%). Studies described participation by 13,640 individuals (median 332; range 31-4601), with sample size data missing from 4% (1/23) of the studies. Across studies, the median percentage of female participants was 77.28% (7404/9581). Sex data were missing from 52% (12/23) of the studies. Age data were variably reported and missing from 57% (13/23) of the studies; therefore, data were not collated. Sex data were missing from 39% (9/23) of the studies. Included studies used a variety of social media platforms to gather research priorities, including websites, emails, Facebook, Twitter, e-newsletters, web-based flyers, Survey Monkey, ExpertLens, blogs, YouTube, Choicebook, Instagram, WhatsApp, Snapchat, and web-based forums. The most common platforms used in the included studies were websites (12/23, 52%) and Facebook (9/23, 39%). The median length of a study's social media campaign, when reported, was 3.5 months (range 1-24 months). Table 1 summarizes the characteristics of the included studies.

Figure 1. Study screening flowchart.

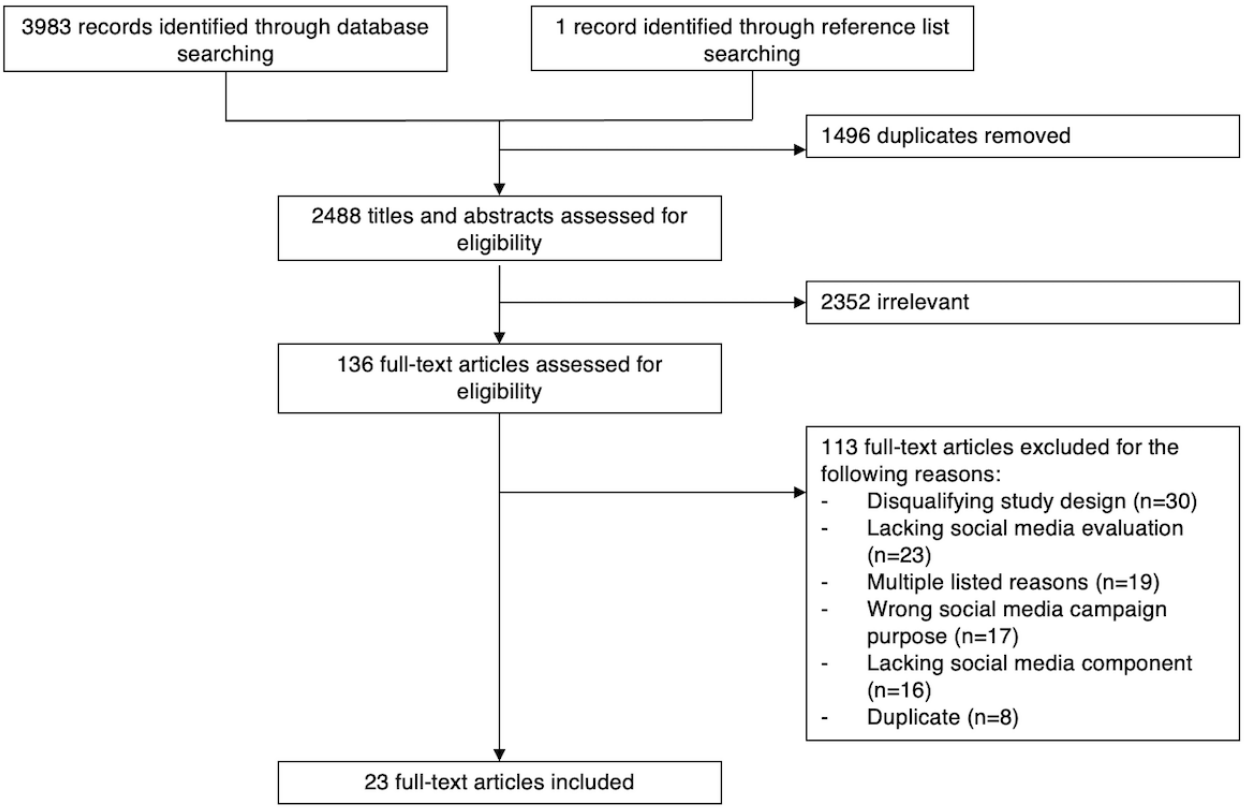


Figure 2. Social media–based research prioritization publication trend.

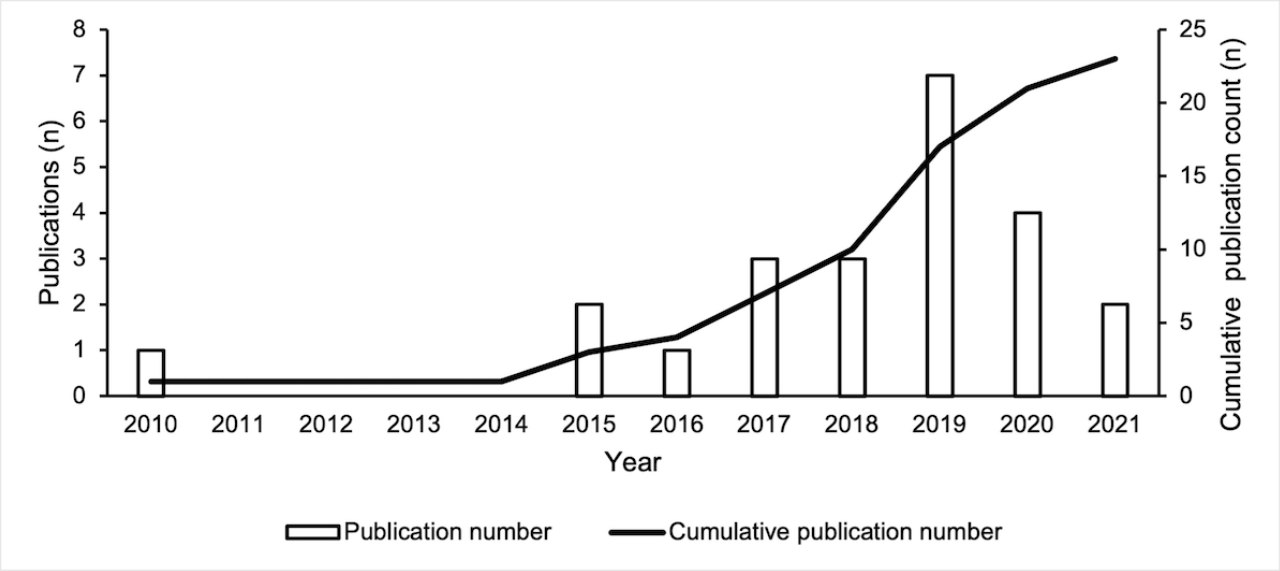


Table 1. Study characteristics (N=23).

Study	Year; country	Sample, N	Age and sex	Social media platform	Social media target group	Purpose for social media use	Duration of social media use	Social media outreach (eg, emails sent and posts made)	Social media analytics (outcomes)	Survey response rate	Outcomes of campaign in terms of research-priority gathering
Allsop et al [12]	2019; 32 countries within Africa	51	Not stated	Website and emails	Members of the African Palliative Care Association and individuals who work in palliative care	To identify (1) current mobile health use in palliative care, (2) potential barriers to use, and (3) priorities for research development	May to August 2016 (4 months)	101 organizations were emailed with web-based survey links	Not stated	51 (100%) survey responses (50.5% response rate)	Research priorities successfully identified
Correll et al [13]	2020; United States	365	Not stated	Website, emails, and other	Patients and caregivers of children (age ≥13 years)	To identify what research topics were most important to patients and caregivers of children with JM[a], JA[b], and cSLE[c]	November 2016, January 2017, and March 2017 for JM, AF[d], and LFA[e], respectively (5 months)	19,176 emails were sent	Not stated	441 survey responses	Research priorities successfully identified
Dyson et al [14]	2017; Canada and Portugal	110	Median age 35 years; 90% (99/110) women, 10% (11/110) men	Website, Facebook, and Twitter	Caregivers of children aged 0-17 years	To identify the outcome priorities of parents of children who had experienced an acute respiratory infection	December 2013 to March 2014 (4 months)	Creation of website, Facebook, and Twitter page or posts with embedded survey links	Website visits (5207); 3.9% view rate	110 (100%) survey respondents	Research priorities successfully identified
Dyson et al [15]	2017; Canada and Portugal	110	Median age 35 years; 90% (99/110) women, 10% (11/110) men	Website, Facebook, and Twitter	Caregivers of children aged 0-17 years	To identify the outcome priorities of parents of children who had experienced an acute respiratory infection	December 2013 to March 2014 (4 months)	Creation of website, Facebook, and Twitter page or posts with embedded survey links	Survey site visits (5027); Facebook page likes (104); and Twitter following (52 new followers)	110 (100%) survey respondents	Research priorities successfully identified

Study	Year; country	Sample, N	Age and sex	Social media platform	Social media target group	Purpose for social media use	Duration of social media use	Social media outreach (eg, emails sent and posts made)	Social media analytics (outcomes)	Survey response rate	Outcomes of campaign in terms of research-priority gathering
Eberman et al [16]	2019; United States	4601; 87 (1.89%) for focus groups, 4514 (98.11%) for survey	Age not stated; 55.05% (2533/4601) women, 43.40% (1997/4601) men, and 0.61% (28/4601) no indication	Newsletters via email	Athletic trainers	To identify research priorities and unify research with clinical practice to improve patient care and advance the profession	January 30, 2017 to March 16, 2017 (2 months)	48,752 emails were sent	Started the survey (5131, 10.5%); agreed to participate (4514, 9.3%); and completed the questionnaire (3910, 86.6%)	4514 (100%) research participants (9.3% response rate)	Research priorities successfully identified
Han et al [17]	2019; United States	332	Median age 51 years; 100% (332/332) women	Newsletters via web, website, Facebook, Twitter, web-based flyers, and emails	Females aged ≥18 years	To identify diabetes type 1 or 2 or prediabetes health research priorities	November 2016 to June 2017 (8 months)	904 website posts	Survey link clicks (421); comments on posts (904); total likes (530); total searches (167); and resource download (671)	332 (100%) research participants	Identified high priority research areas for women living with diabetes
Han et al [18]	2017; United States	332	Median age 49 years; 100% (332/332) women	Newsletters via web, website, Facebook, Twitter, web-based flyers, and emails	Females aged ≥18 years	To identify diabetes type 1 or 2 or prediabetes health research priorities	Not stated	551 emails were sent	Tag clicks (497); reposts and comments (872); voted for posts (540); searched for resources (167); and downloaded resources (671)	332 (100%) survey respondents (84% response rate)	The researchers identified 11 high priority categories of topics that were discussed on the DiabetesSistersVoices community

Study	Year; country	Sample, N	Age and sex	Social media platform	Social media target group	Purpose for social media use	Duration of social media use	Social media outreach (eg, emails sent and posts made)	Social media analytics (outcomes)	Survey response rate	Outcomes of campaign in terms of research-priority gathering
Healy et al [19]	2018; United Kingdom and Ireland	790	Age not stated; 71% (561/790) women, 28.98% (229/790) men	Website, emails, and Twitter	People invited to participate in a randomized trial or participated in Trial Steering Committees, front line randomized trials staff and investigators, and people familiar with trial methodology	To identify priority research questions related to trial recruitment	July 2016 to August 2016 (1 month)	Not stated	Not stated	790 (100%) respondents	List of top 10 trial recruitment uncertainties, determined by those directly involved in trials, were identified
Kim et al [20]	2018; United States	360	Age not stated; 60% (216/360) women, 40% (144/360) men	ExpertLens (ie, expert opinion forums), emails, and other	Patient, patient advocate, clinician, and researcher stakeholders	To determine engagement of stakeholders in research related to heart failure, obesity, and Kawasaki disease	18 months	Not stated	Not stated	84% response rate	Research priority successfully identified
Kriss et al [21]	2019; United States	207	Not stated	Email	Experts in global, regional, and national or subnational health	To identify research priorities for achieving disease elimination goals in the context of measles and rubella	October 17 to November 4, 2016 (approximately 1 month)	774 emails were sent	Not stated	207 (100%) respondents	Four main research priorities within the field of measles and rubella
Morris et al [22]	2015; United Kingdom	475	Not stated	Website, newsletters, and emails with embedded links	Children with neurodisability, caregivers, and clinicians	To identify and prioritize research questions regarding ways to improve the health and well-being of children and young people with neurodisability	Not stated	Creation of website and emails were sent with embedded links	Not stated	369 respondents (78% response rate)	Successfully established top 3 research priorities

Study	Year; country	Sample, N	Age and sex	Social media platform	Social media target group	Purpose for social media use	Duration of social media use	Social media outreach (eg, emails sent and posts made)	Social media analytics (outcomes)	Survey response rate	Outcomes of campaign in terms of research-priority gathering
Morse et al [23]	2021; United States	31	Mean age 15 years; 55% (17/31) women, 45% (14/31) men	Email and social media platforms (not specified)	Parents of children with medical complexity	To (1) ascertain parents' perceived characteristics of child pain experiences, (2) determine the extent to which parents feel that caregivers adequately address pain, and (3) identify ways in which pain collaboration between parents and caregivers may be improved	August 2018 to February 2019 (6 months)	Posting institutional review board–approved message on primary investigator's social media page	Social media post shares (n=30)	Not stated	Established research priorities
Normansell et al [5]	2015; United Kingdom	57	Not stated	Survey Monkey, Facebook, Twitter, website, and other	Patients, caregivers, and health care professionals with expertise in this discipline	To identify research priorities in asthma	August 6 to September 5, 2014 (1 month)	Not stated	"Obtained a large number of responses in a short timeperiod with potentially wide geographical reach"	Not stated	Developed a list of priority Cochrane Reviews
Oesophago-Gastric Anastomosis Study Group [24]	2020; United Kingdom	363	Not stated	WhatsApp and email	OGAA[f] committee, national leaders, and engaged clinicians from high-, low-, and middle-income countries	To prioritize future research areas of unmet clinical need in RCTs[g] to reduce anastomotic leaks	September to November 2019 (3 months)	Posted on organizations' social media accounts	Not stated	Not stated	Established research priorities

Study	Year; country	Sample, N	Age and sex	Social media platform	Social media target group	Purpose for social media use	Duration of social media use	Social media outreach (eg, emails sent and posts made)	Social media analytics (outcomes)	Survey response rate	Outcomes of campaign in terms of research-priority gathering
Rowbotham et al [25]	2019; worldwide	482	Not stated	Twitter	Patients, their caregivers, and clinicians	To identify research priorities for cystic fibrosis	March 2016 to January 2017 (10 months)	320 tweets	Twitter followers gained (n=732); total number of views (n=151,000); engagements with hashtag (n=1806); and followers (n=1160)	Not stated	Top 10 list for research in CF[h] was established
Russell et al [26]	2016; Canada	96	Not stated	Facebook	Family members of children	To exchange knowledge on project planning and research direction and translate research knowledge on disabilities and medical complexity	June 2014 to March 2015 (10 months)	432 Facebook posts were published	96 Facebook members; posts were generally seen by all group members; median likes (n=3); and comments (n=4)	49 respondents (51% response rate)	Provided researchers with an opportunity to consult families of children with special needs to receive guidance and hear issues that are important to them. Research priorities not identified
Salmi et al [27]	2020; United States	36	Not stated	Twitter, emails, blog posts, and Facebook groups	Patients with brain tumor and their care partners (ie, family members and friends who care for patients)	To describe the use of Twitter to complement in-person stakeholder engagement and report emerging themes from qualitative analysis of tweet chats on quality of life needs and palliative care opportunities for patients with brain tumor	April 2018 (1 month)	Two 60-minute scheduled live chat on Twitter	417 tweets by participants in first session and 355 tweets by participants in second session	N/A[i]	Research priorities, in the form of qualitative themes, were successfully identified

Study	Year; country	Sample, N	Age and sex	Social media platform	Social media target group	Purpose for social media use	Duration of social media use	Social media outreach (eg, emails sent and posts made)	Social media analytics (outcomes)	Survey response rate	Outcomes of campaign in terms of research-priority gathering
Shalhub et al [28]	2020; United States, United Kingdom, and Canada	300	Not stated	Blogs and website	Patients and their caregivers	To understand patient needs and determine the research methods best suited to study the adverse health implications associated with vascular Ehlers-Danlos syndrome	January 2018 and April 2018 (2 months)	Not stated	Facebook members in secret group (n=363) and Facebook followers (n=80,573)	Not stated	Optimal modality for research participation and methodologies for building trust in the research teams were identified
Shields et al [29]	2010; Canada	>800	Not stated	Choicebook, message board, blog, YouTube, Facebook, and email	Residents of and health service providers in northwestern Ontario	To engage the disperse population of northwestern Ontario in health care priority setting	Not stated	YouTube video welcome message; weekly blogs; and weekly participation update reports	“Hits” on website platform (n=2500); website views (n=2000); and >800 participants	Not stated	Findings identified new or additional research priorities for health network
Siefried et al [30]	2021; Australia	47	Mean age 42 years; 45% (21/47) women, 45% (21/47) men, and 5% (2/47) other or preferred not to say	Newsletter, emails with embedded links, Twitter, and website	Consumers, family, friends, caregivers, clinicians, researchers, policymakers, industry, research funders, institutions, organizations, law enforcement, border control, and other community members interested in the topic of methamphetamine	To identify clinical research priorities for methamphetamine and emerging drugs of concern in Australia, to guide the work of the National Centre for Clinical Research on Emerging Drugs	February 2019 to March 2019 (1 month)	Newsletter with embedded link were sent to mailing list and recipients of emails were invited to forward the email to other interested parties	Not stated	Not stated	Research themes and priorities were successfully identified

Study	Year; country	Sample, N	Age and sex	Social media platform	Social media target group	Purpose for social media use	Duration of social media use	Social media outreach (eg, emails sent and posts made)	Social media analytics (outcomes)	Survey response rate	Outcomes of campaign in terms of research-priority gathering
Sinclair et al [31]	2019; Croatia, France, Germany, Italy, the Netherlands, Poland, Portugal, Spain, and the United Kingdom	80	Mean age 38 years; 94% (75/80) women, 6% (5/80) men	ConnectE-people (e-forum), Facebook, YouTube, Twitter, WhatsApp, Snapchat, and Instagram	Parents of children with illness	To identify the research priorities of parents of children with Down syndrome, cleft lip or cleft palate, congenital heart defects, or spina bifida	Approximately 2 months	105 parents were invited to secret Facebook group	92% (74/80) of participants accessed the survey through social media and Facebook members (32)	54 (68%) respondents (51.4% response rate)	Top 10 list of research priorities were successfully identified
Sylvia et al [32]	2018; United States	4103	Age range between 18 and 86 years; 78.21% (3209/4103) women, 19.01% (780/4103) men	Website and web-based forums	Patients, caregivers, clinicians, and other advocates	To understand research topics that are of most interest to individuals with mood disorders	May 2015 to May 2017 (24 months)	Not stated	4103 (100%) users enrolled into the web-based community (via the website)	Not stated	Research priority agenda in the area of mood disorders were successfully identified

Study	Year; country	Sample, N	Age and sex	Social media platform	Social media target group	Purpose for social media use	Duration of social media use	Social media outreach (eg, emails sent and posts made)	Social media analytics (outcomes)	Survey response rate	Outcomes of campaign in terms of research-priority gathering
Wojcieszek et al [33]	2019; Australia, New Zealand, Africa, Asia, Europe, North America, South or Central America, the United Kingdom, and Ireland	79	Not stated	Emails with embedded link	Individuals involved in stillbirth research, clinical practice, and advocacy	To identify research priorities and explore potential methodologies to inform care in subsequent pregnancies following a stillbirth	June 2018 to August 2018 (1.5 months)	124 email invitations were sent	Not stated	79 (100%) respondents (64% survey response rate)	Five priority research topics were successfully identified

[a]JM: juvenile myositis.

[b]JA: juvenile arthritis.

[c]cSLE: childhood-onset systemic lupus erythematosus.

[d]AF: Arthritis Foundation.

[e]LFA: Lupus Foundation of America.

[f]OGAA: oesophago-gastric anastomosis audit.

[g]RCT: randomized controlled trial.

[h]CF: cystic fibrosis.

[i]N/A: not applicable.

Research Question 1: Social Media–Based Strategies Used

Table 2 shows the particular social media strategies used to enhance knowledge user engagement in research priority-setting exercises grouped by platform. Of studies using email as their primary social media platform [12,16,17,19,23,24,27,29,30,33] study teams emailed messages with embedded research prioritization survey links (including to researchers' existing mailing lists) and integrated tell a friend tool in emails to prompt recipients to invite colleagues to participate. Facebook-specific methods to engage stakeholders included embedding survey links within Facebook posts, using the platform's boosting feature (ie, paid advertisements), and hiring a Facebook advertising specialist. Informational Facebook pages were also used and involved private and public question-and-answer pages and a resource center with links to relevant documents [5,17,14-18].

Twitter-specific methods to engage participation included the use of hashtags within tweets and question-and-answer threads for prospective participants [5,14,15,17-19,25,31]. In addition, Salmi et al [27], hosted live chats on Twitter, in which host Twitter accounts tweet about predefined topics with questions during a set period, to which Twitter users respond via tweets and engage in discussions with each other. A web-based forum strategy led to the creation of a space where families and researchers could share ideas on the priority-setting research project [31]. Informational videos were created and hosted on YouTube for people potentially interested in contributing research priorities and were later posted on other platforms [28,29]. For studies involving blogs, researchers posted stories and internal updates related to the project to enhance interest in participation [28,29]. Studies also distributed e-newsletters to existing networks, sending them monthly to promote participation [16-18,28,30]. In addition, several studies used posts on Reddit and websites and web-based connection with the research team through video-calling platforms (eg, Skype, WhatsApp, or FaceTime or video chat on Facebook Messenger) to promote participation in priority-setting research [22,28,32].

Table 3 summarizes techniques to disseminate actual web-based research priority-setting surveys through social media. Snowball recruitment, in which current participants' friends and family were approached for participation, was used [14,15,29,30]. Study teams also provided partner organizations with toolkits, templates, and promotional materials [5,12,13,15,17,23,29]. Then, organizations could use these materials to support the

broadcasting of participation opportunities through social media. Individuals embedded in research prioritization exercises, such as steering group members, were additionally asked to promote the participation opportunity to their networks via social media [12-16,19,23,30], including by providing such individuals with preworded statements to tweet [19].

Table 2. Social media platform strategies.

Social media platform and specific strategy	Strategy description	Representative quotes	Studies providing evidence
Blogs			
Blog post stories	Posting insightful stories related to the priority-setting research project with the goal of promoting participation	• "Weekly blogs by the chief executive officer profiling stories that are particularly moving or insightful, as well as internal news on the project."	Shields et al [29]
Project news posting	Posting internal news or updates related to the priority-setting research project	• "Some organisations or individuals promoted the study on Twitter or a blog."	Dyson et al [14]
Emails			
Embedded links	Embedding survey links within emails to promote participation in the priority-setting research project	• "Invitations to participate in the research and a link to the online survey (in the relevant language) were sent via email. Those approached to complete the survey were identified using membership lists of the African Palliative Care Association (APCA)."	Allsop et al [12], Correll et al [13]; Han et al [18], Kriss et al [21], Siefried et al [30], and Wojcieszek et al [33]
Mailing list distribution	The use of an existing mailing list to promote participation in the priority-setting research project	• "A link to an initial electronic survey (created using REDCap) was emailed to members of Cure JM[a], AF[b] and LFA[c] patient and family members and posted on their respective social media sites. The ranking survey was emailed to the Cure JM, AF, and LFA listservs and a link was posted on their respective social media sites."	Allsop et al [12], Correll et al [13], Han et al [17], Siefried et al [30], and Wojcieszek et al [33]
Peer-to-peer dissemination	Using a *tell a friend tool*, which invites friends and colleagues to participate (peer-to-peer messaging) in the priority-setting research project	• "Tell a Friend tool to invite friends or colleagues to participate, using e-mail-based peer-to-peer messaging."	Shields et al [29]
Reminders to participate	Sending email reminders to individuals about the opportunity to participate in the priority-setting research project	"We sent an initial e-mail on Tuesday, January 30, 2017, at 12:00 PM EST to potential participants and, on subsequent Tuesdays between 10:00 AM and 12:00 PM EST, sent 5 weekly reminders to those who had not yet responded."	Eberman et al [16], Han et al [17], Kriss et al [21], and Wojcieszek et al [33]
Reminders to finish survey	Sending email reminders to individuals who began the survey but only partially completed it	• "Reminder emails were sent to non-responders and to individuals who began the survey but only partially completed it."	Kriss et al [21] and Wojcieszek et al [33]
Facebook			

Social media platform and specific strategy	Strategy description	Representative quotes	Studies providing evidence
Embedded links to create ease of participation	Embedding simple and direct links within Facebook posts to external sites related to participation in the priority-setting research project	• "Simple 'How to Participate' area that provided a visual menu of the ways to get involved, with simple links to take participants directly to the tools. Resource Centre page with access to links, documents and reports to help participants deepen their knowledge of the technical health challenges in the region."	Normansell et al [5] and Shields et al [29]
Engagement of advertising strategists	Hiring a Facebook advertising strategist to plan the social media campaign used for promoting participation in the priority-setting research project	• "Tactica Interactive, a digital media enterprise, was hired to broaden our sampling frame via a Facebook advertising strategy."	Dyson et al [15]
Providing participation explanation	Creating a Facebook section that explains how to participate in the priority-setting research project	• "Simple 'How to Participate' area that provided a visual menu of the ways to get involved, with simple links to take participants directly to the tools."	Dyson et al [15]
Use of private and public pages	Creating both public and private Facebook groups to allow private discussion among participants in the priority-setting research project	"Announcement of the vEDS[d] Collaborative survey was disseminated via vEDS public and private social media pages." • "Secret Facebook groups, providing optimal security, were set up for newly recruited research-aware parents (RAPs) to communicate privately and confidentially with each other and for the research team to generate questions and to interpret findings."	Dyson et al, [14], Shalhub et al [28], and Sinclair et al [31]
Providing project explanation	Creating a section on Facebook page dedicated to explaining the priority-setting research project and how participation could have an impact	• "'About our Project' section to provide participants with specific details on how their participation would affect the North West LHIN[e] decision-making and the second IHSP[f]."	Shields et al [29]
Question and answer	Using and moderating a web-based question-and-answer thread on Facebook to promote discussion topics regarding research participation	• "To encourage engagement and re-engagement, the site moderator used online question and answer threads to keep promoting new discussion topics and emailed a weekly topic to all the registered users to encourage them to come back."	Han et al [17] and Sinclair et al [31]

Social media platform and specific strategy		Strategy description	Representative quotes	Studies providing evidence
	Resource center	Creating a resource center with links to documents and reports on the Facebook page	• "'Resource Centre' page with access to links, documents and reports to help participants deepen their knowledge of the technical health challenges in the region."	Shields et al [29]
	Private and secret groups	Creating private Facebook groups to allow private discussion among participants in the priority-setting research project	• "Announcement of the vEDS Collaborative survey was disseminated via vEDS public and private social media pages"	Shalhub et al [28] and Sinclair et al [31]
Newsletter				
	Distribution through the researcher's existing network	Distributing newsletter to an existing network to promote participation in the priority-setting research project	• "To increase our reach and the likelihood of participation, the NATA[g] marketing team distributed our recruitment announcement and link to volunteers via the ''Range of Motion'' newsletter to all registered attendees 5 and 6 weeks before the conference."	Han et al [18], Eberman et al [16], and Siefried et al [30]
	Frequent promotion	Sending monthly newsletters to promote participation in the priority-setting research project	• "Social media promotion through Facebook and Twitter and monthly electronic newsletters from DiabetesSisters."	Han et al [18] and Han et al [17]
Web-based forums	Idea sharing	Creating forums through which families and researchers could share their ideas related to the priority-setting research project	• "Moderated online group where families and researchers can share ideas related to research."	Russell et al [26]
Reddit	Posting of promotional material	The use of Reddit as a social media platform used to promote participation in the priority-setting research project	• "Announcement of the vEDS Collaborative survey was disseminated via vEDS public and private social media pages."	Shalhub et al [28]
Twitter				
	Hashtags	Using Twitter hashtags to attract participants and generate conversation among relevant stakeholders	• "A bespoke Twitter account was set up @questionCF with the associated hashtag #questionCF. This was managed by members of the steering group and aimed to promote the online surveys and increase participation."	Rowbotham et al [25]
	Question and answer	Creating a post for inviting participants to ask questions about the priority-setting research project, which was moderated by steering group members	"A bespoke Twitter account was set up @questionCF with the associated hashtag #questionCF. This was managed by members of the steering group and aimed to promote the online surveys and increase participation."	Rowbotham et al [25]

Social media platform and specific strategy		Strategy description	Representative quotes	Studies providing evidence
	Live chats	Host Twitter accounts tweeting about predefined topics with questions over a set period, during a scheduled chat, to which Twitter users respond via tweets and engage in discussions with each other. Tweets from participants are limited to 280 characters and participants typically include an assigned hashtag in their tweet, thus allowing aggregation of the conversation.	• "The tweet chat hosts (@BTSMchat and @HPMchat, respectively) tweeted the 4 predefined topics (Table 1) with questions over a 60-minute period during a scheduled chat. The hosts alerted tweet chat participants that the transcript of the chat would be subject to qualitative analysis and used to inform research. One tweet question was posted roughly every 15 minutes. Twitter users responded to the questions and engaged in discussions with each other. On Twitter, responses are limited to 280 characters, and participants were instructed to add the #BTSM or #HPM hashtag to aggregate the conversation."	Salmi et al [27]
YouTube	Welcome video	Using YouTube to create a personal welcome message on Facebook pages, inviting users to participate in the priority-setting research project	• "On the site's home page, YouTube video personal welcome message."	Shields et al [29] and Shalhub et al [28]
Website	Posting of promotional material	Discussing the use of websites with survey as a social media platform used to promote participation in the priority-setting research project	• "We created an online and social media presence via a study website (Outcomes in Child Health)..." • "We collaborated with organisations interested in ARI[h] and patient engagement to advertise our research via websites and other channels..."	Allsop et al [12], Dyson et al, Normansell et al [5], and Sylvia et al [32]
Video calling	Digital connection to promote participation	Discussing the use of video-calling or internet-based face-to-face interactions to promote participation in the priority-setting research project	• "Discussed details about the project and the parents' research needs through face-to-face social media platforms such as Skype, WhatsApp, FaceTime, or via video chat on Facebook Messenger to build trust."	Sinclair et al [31]

[a]JM: juvenile myositis.

[b]AF: Arthritis Foundation.

[c]LFA: Lupus Foundation of America.

[d]vEDS: vascular Ehlers-Danlos syndrome.

[e]LHIN: local health integration network.

[f]IHSP: integrated health services plan.

[g]NATA: National Athletic Trainers' Association.

[h]ARI: acute respiratory infection.

Table 3. Dissemination techniques.

Category and specific technique	Technique description	Representative quotes	Studies providing evidence
Existing network			
Individual promotion	Using individuals (eg, steering group members) within existing network to promote the survey to their networks via social media	• "Those approached to complete the survey were identified using membership lists of the African Palliative Care Association (APCA)." • "A link to an initial electronic survey (created using REDCap) was emailed to members of Cure JM, AF and LFA patient and family listservs and posted on their respective social media sites." • "We also asked individuals and organisations within our existing networks to promote the study." • "All Steering Group members were requested to use pre-worded Tweets, which included the link to the survey." • "Invitations to participate in the research and a link to the online survey (in the relevant language) were sent via email. Those approached to complete the survey were identified using membership lists of the African Palliative Care Association (APCA)."	Allsop et al [12], Correll et al [13], Dyson et al, [14], Eberman et al [16], Healy et al [19], Rowbotham et al [25], and Siefried et al [30]
Individual promotion–prewording	Providing individuals (eg, steering group members) within existing network with preworded tweets to promote the research participation opportunity on their Twitter accounts	• "All Steering Group members were requested to use pre-worded Tweets, which included the link to the survey." • "A bespoke Twitter account was set up @questionCF with the associated hashtag #questionCF. This was managed by members of the steering group and aimed to promote the online surveys and increase participation."	Dyson et al [15]; Healy et al [19], Rowbotham et al [25], and Morse et al [23]
External organizations			
Social media collaboration	External organizations posting on their respective social media sites to promote research participation opportunity	• "A link to an initial electronic survey (created using REDCap) was emailed to members of Cure JM, AF and LFA patient or family listservs and posted on their respective social media sites. The ranking survey was emailed to the Cure JM[a], AF[b], and LFA[c] listservs and a link was posted on their respective social media sites." • "Tactica Interactive, a digital media enterprise, was hired to broaden our sampling frame via a Facebook advertising strategy." • "We collaborated with organisations interested in ARI[d] and patient engagement to advertise our research via websites and other channels..." • "A toolkit aimed at partnering organizations, which included a template for the invitation from the partner, a description of DiabetesSistersVoices, and promotional materials including flyers and postcards." • "A survey consisting of 27 questions was developed and distributed to surgeons from the OGAA[e] collaborative and advertised through specialty organizations' social media accounts"	Correll et al [13], Dyson et al [14], Han et al [17], Normansell et al [5], Siefried et al [30], and Oesophago-Gastric Anastomosis Study Group [24]
Providing resources	Providing external organizations with toolkits, templates, or promotional materials that serve as guidelines for when organization broadcasts research participation opportunity	• "A toolkit aimed at partnering organizations, which included a template for the invitation from the partner, a description of DiabetesSistersVoices, and promotional materials including flyers and postcards."	Han et al [17]

Category and specific technique		Technique description	Representative quotes	Studies providing evidence
	Website	External organizations posting on their website to promote research participation opportunity	• "We collaborated with organisations interested in ARI and patient engagement to advertise our research via websites and other channels: The Alberta Centre for Child, Family & Community Research (now known as PolicyWise for Children and Families; a provincial organisation linking government, academia and the community in a focus on evidence-informed policy and practice),22 TRanslating Emergency Knowledge for Kids (a national network of researchers and clinicians invested in improving paediatric emergency care), 23 the Cochrane Consumer Network (an international network of healthcare consumers with an interest in evidence-based medicine) 24 and the Stollery Family Centered Care Network (a local children's hospital-based network of patients and families that provide input into patient care)." • "Online survey was posted on Survey Monkey and advertised through the Asthma UK Facebook and Twitter profiles and Cochrane Airways social media and website."	Allsop et al [12], Dyson et al [14], and Normansell et al [5]
Snowball recruitment	N/A[f]	Disseminating research opportunity to participants' social networks to increase participation and access to specific populations	• "We used snowball sampling to recruit parents." • "First, we focused on identifying and engaging recruitment targets with the potential for a high yield of participants. We then expanded our scope through referrals and diffusion via social media." • "Through Facebook, friend networks were encouraged to invite each other to participate." • "Tell a Friend tool to invite friends or colleagues to participate, using e-mail-based peer-to-peer messaging."	Dyson et al [14], Shields et al [29], and Siefried et al [30]
Boosts	N/A	Using the Facebook *boosting* feature to reach a wider audience of possible participants	• "Facebook posts were "boosted" monthly to showcase the posts to more users." • "Social media promotion through Facebook and Twitter and monthly e-newsletters from DiabetesSisters Facebook posts were boosted to showcase the posts to more users, centralizing it to female users in the United States with interests in diabetes-relevant topics. DiabetesSisters posted on Facebook about the study and each month they "boosted" the post to increase the number of women who saw each post."	Han et al [17] and Han et al [18]

[a]JM: juvenile myositis.

[b]AF: Arthitis Foundation.

[c]LFA: Lupus Foundation of America.

[d]ARI: acute respiratory infection.

[e]OGAA: oesophago-gastric anastomosis audit.

[f]N/A: not applicable.

Research Question 2: Measurement of Social Media Campaign Effectiveness

Across all the 23 included studies, 21 (91%) claimed to be successful in conducting health research priority-setting exercises via social media–based methods.

The direct effect of social media campaigns in securing stakeholder participation in research priority-setting was assessed as the (1) number of survey responses [12,14,15,20,33], (2) number of survey responses within a set period [14,15,20,33], (3) proportion of surveys fully completed [21], and (4) number of visits to external survey administration sites [14,15].

Indirect metrics for campaign effectiveness were (1) audience reach (ie, extent to which the survey sample was characteristic of the target population [13-15], number of countries and local communities represented in the sample [12,21], and number of national associations and external organizations contacted [12]);

(2) campaign interaction (ie, number of clicks and impressions on posts [14,15,18,23,25,27], frequency of post views [26], volume of comments left by target stakeholders [26], number of searches for campaign pages or downloads of resources [17,18], number of bespoke hashtag clicks or uses [25,27], and Google Analytics [18]); (3) participant satisfaction [17,28,31]; and (4) platform-specific methods (ie, number of website views or likes [12,14,15,17-19,21,29], number of registered participants in an email chain or total number of delivered emails [12,13,16,19,21,29,33], new followers and likes on Facebook pages [14,15,17,18,26], and Twitter followers gained [14,15,25]).

Research Question 3: Benefits, Limitations, and Recommendations

Benefits and Limitations of Social Media–Based Research Priority Setting

All included studies (23/23, 100%) successfully gathered research priorities from key stakeholders and knowledge users using social media–based participant recruitment. Cited benefits related to social media use were the capacity to elicit participation from many knowledge users [14,15,17,18,27,31], the speed at which research priorities were gathered, the sense of community developed [17,31], peer-support offered to patients and family members [17,26,28,31] by social media campaigns, and the capacity for dissemination of health-promoting resources from health care professionals to patients. A cited limitation of social media–based methods was that web-only methods may limit the participation of individuals with limited or no access to technology, limited leisure time to engage with social media, and lower socioeconomic status and of older age [12-15,17].

Recommendations for Successful Social Media–Based Research Priority Setting

To improve the effectiveness of social media campaigns, authors recommended focusing on the campaign's graphic design components and style of messaging [26,31,32], creating opportunities for the target audience to personally interact with the team leading the campaign [31], and using platform-specific paid advertisements (ie, also termed *boosts*) [18,28].

Design-related recommendations included implementing illustrative and graphical sophistication, such as posts containing words, text, and video [31] and establishing a tone and style of graphics to create a consistent brand [26,32]. Messaging recommendations were to post some content that is not directly related to research, but of interest to community members—especially if these posts are community-led [22,26,31]; to avoid phrases that do not foster inclusivity and may separate the researchers from the target audience (ie, *us* vs *them* semantics); and to minimize scientific jargon in posts. Interaction-related recommendations involved using moderators [17,26], especially community members to build the authenticity of the campaign [27]; initiating conversations with perspective participants to *break the ice*; using software that supports face-to-face interaction between researchers and the community [31]; allowing peer-to-peer sharing (ie, providing community members with capacity to invite colleagues to participate) [17,22,26,28,29,31,33]; and using platform-specific boosts (eg, Facebook boosts) [18,28]. This last strategy corresponded with the highest recruitment and enrollment yields.

Recommendations to address the limitation that social media may prevent priority-setting participation by some groups were also suggested. These included implementing a hybrid of electronic and nonelectronic survey dissemination methods to increase the representation of those without access to technology [12,17,18], developing web-based materials with simple navigation requirements to allow participation by individuals with less experience with the web [30], and intentionally tailoring social media strategies (eg, hashtags and boosts) for subpopulations of individuals whom study teams identify as being underrepresented in research prioritization project data sets [13-15,17,21,25,32].

Discussion

Principal Findings

Recognizing the importance of engaging key stakeholders in developing research agendas, we sought to use the extant literature to understand how social media might support research priority-setting, how effectiveness of the method might be measured, and the method's benefits and drawbacks. We show that multiple social media strategies, which differ depending on the social media platform, have been used to promote participation in research priority setting—with strong success rates in generating research agendas. Metrics to quantify the reach of these strategies included the number of impressions on posts (eg, likes and other reactions) and the volume of comments left by stakeholders. In addition to the benefits, limitations of the use of social media in research priority-setting were also identified. Results from this review can guide methods for research priority-setting by patients, family caregivers, health care professionals, and other advocates and support the engagement of these stakeholders in developing future research agendas.

Social Media Platform Strategies and Dissemination Techniques

Social media–based strategies that incorporated platform-specific amplification (eg, Facebook boosts) and components that encouraged active engagement by participants (eg, question-and-answer forums and shared resources) enabled researchers to reach a broad audience of possible participants. This finding agrees with the literature showing that Facebook [34] health promotion posts receiving a paid boost reached significantly more users. Hashtags were also used in the included studies to increase visibility of tweets, which aligns with previous research showing hashtag use as effective in influencing social media conversations related to mental health [35] and in cases where the desired participant pool is small [36].

Our finding that snowball sampling is used to disseminate priority-setting surveys and expand participant pools aligns with other research showing that options to like, tag, or share posts expand a social media campaign's reach [37]. This method may be particularly advantageous in cases where the campaign target

audience is a specific and relatively small group (eg, people with lived experiences of less common diseases) and campaign participants may have contacts within their social network who they can engage in the process. Our results also suggest that there are priority-setting advantages in asking relevant external organizations and internal research and clinical team members to circulate survey links and use their personal or organization-affiliated social media accounts to expand reach.

Measurement of Social Media Campaign Effectiveness

We identified several metrics used by researchers to evaluate the effectiveness of social media campaigns, including the number of post impressions, frequency of viewed posts, volume of comments left by stakeholders, and number of times a bespoke hashtag was clicked or used. The heterogeneity in metrics likely reflects the exponentially growing number of social media platforms. However, the collection and interpretation of these social media impact metrics support ongoing consideration of the campaign's effectiveness and subsequent content adjustments to maximize campaign reach and engagement [35].

Benefits and Limitations of and Recommendations for Social Media Campaigns in Research

Commonly identified benefits of priority-setting via social media include the speed at which participation opportunities can be disseminated and the capacity to build a sense of community among participants—possibly enhancing engagement. Research has also indicated that social media may be particularly useful in targeting information at some rarely reached groups such as individuals with depression [38]. In addition, moderators might humanize the campaign, build possible participant's trust, and enhance campaign engagement by these individuals [39].

In contrast, limitations of social media–based methods for priority-setting research include the uncertainty of who is being captured through the posts [40]. Our study found that researchers commonly cite fears that social media–based methods may unexpectedly include or exclude the research priority perspectives of certain groups. In these cases, there are limited ways to assure that the recruited team of participants is the valid group of people that will render reliable results. This is problematic from ethical and methodological (ie, sampling bias) points of view and its mitigation requires careful planning. Moreover, when survey links are disseminated via social media, the true number of individuals that are reached cannot be calculated. This is because not all users will engage (ie, like, comment, and share) with the post [13,20]. In addition, although the platform analytics (ie, number of follows, comments, and likes on posts) are often used as an indication of survey engagement, these data may not be representative of the sample that opens the survey link or completes the survey.

Recommendations were also made to establish a consistent tone and branding, with a focus on using attractive graphic designs within priority-setting research campaigns. This consistency may increase the recognizability of the research project and authenticity to the effort, resulting in increased participation in priority-setting research efforts [41].

Limitations of Our Study

The definition of social media varies substantially in the literature and some definitions used did not meet our inclusion criteria. Our conclusions regarding the recruitment for priority-setting research projects may differ from those arising if a different definition was used. Varying definitions of *social media* may also have rendered our decision-making process during the screening phase susceptible to error. However, we screened in duplicate with good consistency and used third-party arbitration of discrepancies. Finally, amid the COVID-19 pandemic, the number of studies adapting to web-based research methodologies, especially using social media, may have increased after the search strategy was performed. Considering such rapid growth, it is important to note that this review is a snapshot at a particular point in time that does not account for novel methods that may have emerged after our search.

Recommendations for Practice and Future Research

Social media appears to be an effective means to recruit and involve participants in the research process. Thus, researchers should consider using web-based social networking as a method to recruit knowledge users, collect data, and translate knowledge into practice. The study team's efforts to build knowledge user trust in prioritization efforts, including by humanizing the campaign through moderating chats and engaging with participants, may improve engagement. On the basis of our findings, efforts can be supported by optimizing the visual representation of data through illustrative posts containing text and graphics. Moreover, to enhance participation by a wide group of knowledge users, researchers should focus on developing accessible and inclusive web-based materials. In addition, investing in platform-specific boosts (eg, Facebook boosts) and paid advertisements may be an effective tactic to enhance participant recruitment and enrollment.

Given the relatively recent emergence of digital platforms, social media–based methods are understudied compared with traditional recruitment means. We have identified some possible limitations of the method, such as potential limited access to individuals of lower socioeconomic status or older age. However, few studies have determined the extent to which these limitations impact prioritization efforts and, in the case of older adults, contrary evidence exists indicating good engagement with social media and technologies [42]. Should the identified limitations of social media–based priority-setting be significant, research into ways to mitigate these shortcomings is needed. Further research is needed to understand how to enhance the capacity of social media recruitment to capture representative samples. More research is also needed to understand which social media strategies and dissemination techniques are likely to be successful for research prioritization efforts, with the understanding that these strategies and techniques are likely to change over time as new social media platforms and features become available. Finally, given the highly public nature of information exchange on social media, considerations of the data privacy and security implications of social media–based research prioritization efforts are needed.

Conclusions

In this review, we synthesized the rapidly emerging data assessing the effectiveness of social media strategies to engage knowledge users in research priority-setting efforts across several social media platforms. The benefits of social media–based recruitment included the speed at which participation opportunities can be disseminated and the sense of community built among participants. As it is likely that social media–based research methods, including for research priority-setting, will be increasingly used by the scientific community, lessons and recommendations from this review can support scientists to more fully engage those who are most impacted by health research in setting associated research agendas.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File , 21 KB - jmir_v24i2e29821_app1.docx]

References

1. Jibb LA, Stacey D, Carley M, Davis A, Graham ID, Green E, et al. Research priorities for the pan-Canadian oncology symptom triage and remote support practice guides: a modified nominal group consensus. Curr Oncol 2019;26(3):173-182 [FREE Full text] [doi: 10.3747/co.26.4247] [Medline: 31285662]
2. Jull J, Giles A, Graham ID. Community-based participatory research and integrated knowledge translation: advancing the co-creation of knowledge. Implement Sci 2017;12(1):150 [FREE Full text] [doi: 10.1186/s13012-017-0696-3] [Medline: 29258551]
3. Harding K, Aryeetey R, Carroll G, Lasisi O, Pérez-Escamilla R, Young M. Breastfeed4Ghana: design and evaluation of an innovative social media campaign. Matern Child Nutr 2020;16(2):e12909 [FREE Full text] [doi: 10.1111/mcn.12909] [Medline: 31867865]
4. Jawad M, Abass J, Hariri A, Akl EA. Social media use for public health campaigning in a low resource setting: the case of waterpipe tobacco smoking. Biomed Res Int 2015;2015:562586 [FREE Full text] [doi: 10.1155/2015/562586] [Medline: 26273631]
5. Normansell R, Welsh E. "Asthma can take over your life but having the right support makes that easier to deal with." Informing research priorities by exploring the barriers and facilitators to asthma control: a qualitative analysis of survey data. Asthma Res Pract 2015;1:11 [FREE Full text] [doi: 10.1186/s40733-015-0011-5] [Medline: 27965765]
6. Snelson CL. Qualitative and mixed methods social media research: a review of the literature. Int J Qual Methods 2016;15(1):160940691562457. [doi: 10.1177/1609406915624574]
7. Kaushal A, Kassianos AP, Sheringham J, Waller J, von Wagner C. Use of social media for cancer prevention and early diagnosis: scoping review protocol. BMJ Open 2020;10(2):e033592 [FREE Full text] [doi: 10.1136/bmjopen-2019-033592] [Medline: 32102815]
8. McGowan J, Straus S, Moher D, Langlois EV, O'Brien KK, Horsley T, et al. Reporting scoping reviews-PRISMA ScR extension. J Clin Epidemiol 2020;123:177-179. [doi: 10.1016/j.jclinepi.2020.03.016] [Medline: 32229248]
9. Namkoong K, Nah S, Record RA, Van Stee SK. Communication, reasoning, and planned behaviors: unveiling the effect of interactive communication in an anti-smoking social media campaign. Health Commun 2017;32(1):41-50. [doi: 10.1080/10410236.2015.1099501] [Medline: 27119592]
10. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
11. Thorne S. Data analysis in qualitative research. Evid Based Nurs 2000;3(3):68-70. [doi: 10.1136/ebn.3.3.68] [Medline: 28850992]
12. Allsop MJ, Namisango E, Powell RA. A survey of mobile phone use in the provision of palliative care services in the African region and priorities for future development. J Telemed Telecare 2019;25(4):230-240. [doi: 10.1177/1357633X18767187] [Medline: 29635966]
13. Correll CK, Dave M, Paul AF, Gaizo VD, Schrandt S, Partovi RS, et al. Identifying research priorities among patients and families of children with rheumatic diseases living in the United States. J Rheumatol 2020;47(12):1800-1806. [doi: 10.3899/jrheum.190934] [Medline: 32062607]
14. Dyson MP, Shave K, Gates A, Fernandes RM, Scott SD, Hartling L. Which outcomes are important to patients and families who have experienced paediatric acute respiratory illness? Findings from a mixed methods sequential exploratory study. BMJ Open 2017;7(12):e018199 [FREE Full text] [doi: 10.1136/bmjopen-2017-018199] [Medline: 29247097]
15. Dyson MP, Shave K, Fernandes RM, Scott SD, Hartling L. Outcomes in child health: exploring the use of social media to engage parents in patient-centered outcomes research. J Med Internet Res 2017;19(3):e78 [FREE Full text] [doi: 10.2196/jmir.6655] [Medline: 28302593]

16. Eberman LE, Walker SE, Floyd RT, Covassin T, Nolton E, Valier AR, et al. The prioritized research agenda for the athletic training profession: a report from the strategic alliance research agenda task force. J Athl Train 2019;54(3):237-244 [FREE Full text] [doi: 10.4085/1062-6050-374-18] [Medline: 30870008]
17. Han P, Nicholson W, Norton A, Graffeo K, Singerman R, King S, et al. DiabetesSistersVoices: virtual patient community to identify research priorities for women living with diabetes. J Med Internet Res 2019;21(5):e13312 [FREE Full text] [doi: 10.2196/13312] [Medline: 31094360]
18. Han P, Nicholson W, Norton A, Singerman A, Sunderesan A, Bennett WL. Development and preliminary findings of diabetes sisters voices-an online community to engage women with diabetes about research and healthcare priorities. J Gen Intern Med 2017;32(2):S159 [FREE Full text]
19. Healy P, Galvin S, Williamson PR, Treweek S, Whiting C, Maeso B, et al. Identifying trial recruitment uncertainties using a James Lind Alliance Priority Setting Partnership - the PRioRiTy (Prioritising Recruitment in Randomised Trials) study. Trials 2018;19(1):147 [FREE Full text] [doi: 10.1186/s13063-018-2544-4] [Medline: 29490702]
20. Kim KK, Khodyakov D, Marie K, Taras H, Meeker D, Campos HO, et al. A novel stakeholder engagement approach for patient-centered outcomes research. Med Care 2018;56(10 Suppl 1):S41-S47 [FREE Full text] [doi: 10.1097/MLR.0000000000000790] [Medline: 30074950]
21. Kriss JL, Grant GB, Moss WJ, Durrheim DN, Shefer A, Rota PA, et al. Research priorities for accelerating progress toward measles and rubella elimination identified by a cross-sectional web-based survey. Vaccine 2019;37(38):5745-5753 [FREE Full text] [doi: 10.1016/j.vaccine.2019.02.058] [Medline: 30898393]
22. Morris C, Simkiss D, Busk M, Morris M, Allard A, Denness J, et al. Setting research priorities to improve the health of children and young people with neurodisability: a British Academy of Childhood Disability-James Lind Alliance Research Priority Setting Partnership. BMJ Open 2015;5(1):e006233 [FREE Full text] [doi: 10.1136/bmjopen-2014-006233] [Medline: 25631309]
23. Morse BL, Serna RW, LaLumiere M, Rogal M, Foley K, Lombardo M, et al. Leveraging parent pain perspectives to improve pain practices for children with medical complexity. Pain Manag Nurs 2021;22(2):169-176. [doi: 10.1016/j.pmn.2020.11.011] [Medline: 33384241]
24. Oesophago-Gastric Anastomosis Study Group on the West Midlands Research Collaborative. Anastomotic leak following oesophagectomy: research priorities from an international Delphi consensus study. Br J Surg 2021;108(1):66-73. [doi: 10.1093/bjs/znaa034] [Medline: 33640931]
25. Rowbotham NJ, Smith SJ, Elliott ZC, Leighton PA, Rayner OC, Morley R, et al. Adapting the James Lind Alliance priority setting process to better support patient participation: an example from cystic fibrosis. Res Involv Engagem 2019;5:24 [FREE Full text] [doi: 10.1186/s40900-019-0159-x] [Medline: 31452934]
26. Russell DJ, Sprung J, McCauley D, Kraus de Camargo O, Buchanan F, Gulko R, et al. Knowledge exchange and discovery in the age of social media: the journey from inception to establishment of a parent-led Web-based research advisory community for childhood disability. J Med Internet Res 2016;18(11):e293 [FREE Full text] [doi: 10.2196/jmir.5994] [Medline: 27836818]
27. Salmi L, Lum HD, Hayden A, Reblin M, Otis-Green S, Venechuk G, et al. Stakeholder engagement in research on quality of life and palliative care for brain tumors: a qualitative analysis of #BTSM and #HPM tweet chats. Neurooncol Pract 2020;7(6):676-684 [FREE Full text] [doi: 10.1093/nop/npaa043] [Medline: 33304602]
28. Shalhub S, Sage L, Demasi J, Wallace SE, Fulton DS, Bloom L, Vascular Ehlers-Danlos Syndrome Collaborative. Assessment of the information sources and interest in research collaboration among individuals with Vascular Ehlers-Danlos Syndrome. Ann Vasc Surg 2020;62:326-334. [doi: 10.1016/j.avsg.2019.06.010] [Medline: 31449940]
29. Shields K, DuBois-Wing G, Westwood E. Share your story, shape your care: engaging the diverse and disperse population of Northwestern Ontario in healthcare priority setting. Healthc Q 2010;13(3):86-90. [doi: 10.12927/hcq.2010.21821] [Medline: 20523159]
30. Siefried KJ, Ezard N, Christmass M, Haber P, Ali R, NCCRED Methamphetamine And Emerging Drugs Clinical Research Network Working Group. A clinical research priority setting study for issues related to the use of methamphetamine and emerging drugs of concern in Australia. Drug Alcohol Rev 2022;41(2):309-319. [doi: 10.1111/dar.13350] [Medline: 34237176]
31. Sinclair M, McCullough JE, Elliott D, Latos-Bielenska A, Braz P, Cavero-Carbonell C, et al. Exploring research priorities of parents who have children with down syndrome, cleft lip with or without cleft palate, congenital heart defects, or spina bifida using ConnectEpeople: a social media coproduction research study. J Med Internet Res 2019;21(11):e15847 [FREE Full text] [doi: 10.2196/15847] [Medline: 31763986]
32. Sylvia L, Hearing CM, Montana RE, Gold AK, Walsh SL, Janos JA, et al. MoodNetwork: an innovative approach to patient-centered research. Med Care 2018;56(10 Suppl 1):S48-S52 [FREE Full text] [doi: 10.1097/MLR.0000000000000789] [Medline: 30074951]
33. Wojcieszek AM, Heazell AE, Middleton P, Ellwood D, Silver RM, Flenady V. Research priorities and potential methodologies to inform care in subsequent pregnancies following stillbirth: a web-based survey of healthcare professionals, researchers and advocates. BMJ Open 2019;9(6):e028735 [FREE Full text] [doi: 10.1136/bmjopen-2018-028735] [Medline: 31230027]

34. Kite J, Grunseit A, Li V, Vineburg J, Berton N, Bauman A, et al. Generating engagement on the make healthy normal campaign Facebook page: analysis of Facebook analytics. JMIR Public Health Surveill 2019;5(1):e11132 [FREE Full text] [doi: 10.2196/11132] [Medline: 31344679]
35. Schlichthorst M, King K, Turnure J, Sukunesan S, Phelps A, Pirkis J. Influencing the conversation about masculinity and suicide: evaluation of the man up multimedia campaign using Twitter data. JMIR Ment Health 2018;5(1):e14 [FREE Full text] [doi: 10.2196/mental.9120] [Medline: 29449203]
36. Rocha HM, Savatt JM, Riggs ER, Wagner JK, Faucett WA, Martin CL. Incorporating social media into your support tool box: points to consider from genetics-based communities. J Genet Couns 2018;27(2):470-480 [FREE Full text] [doi: 10.1007/s10897-017-0170-z] [Medline: 29130143]
37. Bennetts SK, Hokke S, Crawford S, Hackworth NJ, Leach LS, Nguyen C, et al. Using paid and free Facebook methods to recruit Australian parents to an online survey: an evaluation. J Med Internet Res 2019;21(3):e11206 [FREE Full text] [doi: 10.2196/11206] [Medline: 30839282]
38. Clarke J, van Amerom G. A comparison of blogs by depressed men and women. Issues Ment Health Nurs 2008;29(3):243-264. [doi: 10.1080/01612840701869403] [Medline: 18340611]
39. Urbanoski K, van Mierlo T, Cunningham J. Investigating patterns of participation in an online support group for problem drinking: a social network analysis. Int J Behav Med 2017;24(5):703-712. [doi: 10.1007/s12529-016-9591-6] [Medline: 27549786]
40. Pierce M, McManus S, Jessop C, John A, Hotopf M, Ford T, et al. Says who? The significance of sampling in mental health surveys during COVID-19. Lancet Psychiatry 2020;7(7):567-568 [FREE Full text] [doi: 10.1016/S2215-0366(20)30237-6] [Medline: 32502467]
41. Miles RC, Patel AK. The radiology Twitterverse: a starter's guide to utilization and success. J Am Coll Radiol 2019;16(9 Pt A):1225-1231. [doi: 10.1016/j.jacr.2019.03.014] [Medline: 31092350]
42. Anderson M, Perrin A. Tech adoption climbs among older adults. Pew Research Center: internet technology. 2017. URL: https://www.pewresearch.org/internet/2017/05/17/technology-use-among-seniors/ [accessed 2021-10-08]

Abbreviations

PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by A Mavragani; submitted 05.05.21; peer-reviewed by M Gottlieb, MDG Pimentel, M Sinclair, R Halkes; comments to author 30.07.21; revised version received 25.11.21; accepted 23.12.21; published 21.02.22.

Please cite as:
Sivaratnam S, Hwang K, Chee-A-Tow A, Ren L, Fang G, Jibb L
Using Social Media to Engage Knowledge Users in Health Research Priority Setting: Scoping Review
J Med Internet Res 2022;24(2):e29821
URL: https://www.jmir.org/2022/2/e29821
doi:10.2196/29821
PMID:35188476

Review

Digitalized Cognitive Behavioral Interventions for Depressive Symptoms During Pregnancy: Systematic Review

Wan Mohd Azam Wan Mohd Yunus[1,2,3], MClinPsych, PhD; Hanna-Maria Matinolli[1,2], MHealthSci, PhD; Otto Waris[1,2,4], PhD; Subina Upadhyaya[1,2], PhD; Miika Vuori[1,2,5], RMN, PhD; Tarja Korpilahti-Leino[1,2], MA; Terja Ristkari[1,2], MNSc; Tarja Koffert[1], CCPS, RMN; Andre Sourander[1,2,6], MD, PhD

[1]Research Centre for Child Psychiatry, University of Turku, Turku, Finland

[2]INVEST Research Flagship, University of Turku, Turku, Finland

[3]Department of Psychology, School of Human Resource Development & Psychology, Faculty of Social Sciences & Humanities, Universiti Teknologi Malaysia, Johor, Malaysia

[4]Turku University of Applied Sciences, Turku, Finland

[5]Department of Teacher Education, Turku Institute for Advanced Studies, University of Turku, Turku, Finland

[6]Turku University Hospital, Turku, Finland

Corresponding Author:
Andre Sourander, MD, PhD
Research Centre for Child Psychiatry
University of Turku
Lemminkäisenkatu 3/Teutori 3. floor
Turku, 20014
Finland
Phone: 358 503653447
Email: andsou@utu.fi

Abstract

Background: Studies have shown a high prevalence of depression during pregnancy, and there is also evidence that cognitive behavioral therapy (CBT) is one of the most effective psychosocial interventions. Emerging evidence from randomized controlled trials (RCTs) has shown that technology has been successfully harnessed to provide CBT interventions for other populations. However, very few studies have focused on their use during pregnancy. This approach has become increasingly important in many clinical areas due to the COVID-19 pandemic, and our study aimed to expand the knowledge in this particular clinical area.

Objective: Our systematic review aimed to bring together the available research-based evidence on digitalized CBT interventions for depression symptoms during pregnancy.

Methods: A systematic review of the Web of Science, Cochrane Central Register of Controlled Trials, CINAHL, MEDLINE, Embase, PsycINFO, Scopus, ClinicalTrials.gov, and EBSCO Open Dissertations databases was carried out from the earliest available evidence to October 27, 2021. Only RCT studies published in English were considered. The PRISMA (Preferred Reporting Items of Systematic Reviews and Meta-analyses) guidelines were followed, and the protocol was registered on the Prospective Register of Systematic Reviews. The risk of bias was assessed using the revised Cochrane risk-of-bias tool for randomized trials.

Results: The review identified 7 studies from 5 countries (the United States, China, Australia, Norway, and Sweden) published from 2015 to 2021. The sample sizes ranged from 25 to 1342 participants. The interventions used various technological elements, including text, images, videos, games, interactive features, and peer group discussions. They comprised 2 guided and 5 unguided approaches. Using digitalized CBT interventions for depression during pregnancy showed promising efficacy, with guided intervention showing higher effect sizes (Hedges g=1.21) than the unguided interventions (Hedges g=0.14-0.99). The acceptability of the digitalized CBT interventions was highly encouraging, based on user feedback. Attrition rates were low for the guided intervention (4.5%) but high for the unguided interventions (22.1%-46.5%). A high overall risk of bias was present for 6 of the 7 studies.

Conclusions: Our search only identified a small number of digitalized CBT interventions for pregnant women, despite the potential of this approach. These showed promising evidence when it came to efficacy and positive outcomes for depression

symptoms, and user feedback was positive. However, the overall risk of bias suggests that the efficacy of the interventions needs to be interpreted with caution. Future studies need to consider how to mitigate these sources of biases. Digitalized CBT interventions can provide prompt, effective, evidence-based interventions for pregnant women. This review increases our understanding of the importance of digitalized interventions during pregnancy, including during the COVID-19 pandemic.

Trial Registration: PROSPERO International Prospective Register of Systematic Reviews CRD42020216159; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=216159

(J Med Internet Res 2022;24(2):e33337) doi:10.2196/33337

KEYWORDS

pregnancy; antenatal depression; systematic review; cognitive behavior therapy; digital interventions; COVID-19

Introduction

It is estimated that one-fifth of mothers experience depression during pregnancy [1]. This has been strongly associated with postnatal depression [2,3] and linked with several adverse health and psychosocial outcomes for both the mother and child. These include preterm delivery [4,5], difficulties in mother-child interactions, impaired cognitive and psychomotor development [6], and altered fetal brain development [7,8]. Pregnant women experience changes to psychological symptoms due to complex biological and psychological interactions throughout the pregnancy period [9]. Therefore, recognizing depression symptoms during pregnancy at an early stage and preventing future risk of depression should be considered an important global public health challenge.

There is evidence that psychosocial treatment, cognitive behavioral therapy (CBT), and interpersonal psychotherapy are effective in treating depression during pregnancy [10]. The US Preventive Services Task Force [11] reviewed pharmacological and nonpharmacological interventions for depression during pregnancy and postnatal periods. It reported that CBT was effective and stated that global studies had not reported harmful outcomes for either mothers or infants. A randomized controlled trial (RCT) of 217 pregnant women in the United States showed that CBT significantly reduced depression during pregnancy when it was compared with standard care [12]. Concerns about the safety of using selective serotonin reuptake inhibitors during pregnancy and their effect on offspring have been reported, such as altering brain circuits [13], increasing the risk for later depression [14], and increased problems with motor and language development [15]. These highlight the need to study psychosocial interventions.

There is huge gap between what pregnant women need and the human resources that are available to provide such services. Very few women seek help with mental health issues during pregnancy [16]. Barriers to seeking help have included individual factors, such as lack of motivation, time constraints, or they just decided not to seek care [16,17]. Social factors have included stigma and lack of social support, and practical issues have included lack of money or childcare. There are also structural barriers, like lack of information about services, treatment provider responsivity, and treatment accessibility issues [16-20]. The World Health Organization's Mental Health Atlas also reported unequal and limited availability of mental health services and resources across the world [21]. Depression is common during pregnancy, and it needs to be treated without delay. The Institute of Medicine recognized the different spectrum of interventions for mental health based on the risk for the target population: prevention (universal, selective, indicated), treatment (case identification, standard treatment for known disorders), and maintenance (compliance with long-term treatment, aftercare) [22]. Focusing on pregnant women with different depression risk levels is essential, given that less than three-quarters of pregnant women are screened for depression [23].

Digitalized interventions provide an effective way of reaching people and have the potential to overcome barriers, such as the fear of being stigmatized, practical reasons, and treatment availability [24-28]. Digitalized interventions are also becoming increasingly important, as mental health services struggle to deal with the considerable increase in demand for services due to the COVID-19 pandemic. This health emergency began at a time when resources were already scarce, and the number of referrals is now increasing. The use of digital interventions has increased during the COVID-19 pandemic, especially during physical distancing measures, and this has highlighted the increased role it can play in self-care and remote care. Digital interventions have the potential to improve the health outcomes of pregnant women and their experiences with care [29].

To our knowledge, no review has specifically focused on digitalized CBT interventions for depression symptoms during pregnancy. We note that 4 broader reviews on psychosocial interventions have been conducted in the past 5 years, and these can be divided into 2 categories. One review focused on any psychosocial interventions during pregnancy [30], and 3 reviews focused on psychosocial interventions throughout the broad perinatal period, namely pregnancy and the postnatal period [31-33]. The aim of the study by Li et al [30] was to carry out a systematic review and meta-analysis of any psychotherapies for pregnant women that focused on depression, anxiety, and quality of life. However, the authors only identified 2 RCTs on digitalized CBT. It was notable that their review did not include important search terms widely used for digital or internet interventions, such as web-based interventions, digital health, internet, and online interventions [34]. The other 3 reviews mainly identified interventions during the postnatal period. Lee et al [32] did not find any digitalized CBT interventions during pregnancy, while the other 2 studies [31,33] only found limited evidence. It is important to note that none of those 4 broader reviews contained extensive discussions of digitalized CBT interventions.

We decided that a new review was warranted, given the broad nature of previous systematic reviews and the increasing need for, and emergence of, digitalized CBT interventions for depression symptoms during pregnancy. Our systematic review aimed to evaluate the efficacy and acceptability of digitalized CBT interventions for depression symptoms during pregnancy. This review had 2 specific aims. The first was to evaluate the efficacy of digitalized CBT interventions during pregnancy by comparing the effect sizes between studies. The second was to assess the acceptability of digitalized CBT interventions by looking at attrition rates and feedback from participants.

Methods

This systematic review was conducted in accordance with the PRISMA (Preferred Reporting Items of Systematic Reviews and Meta-analyses) [35] and the Synthesis Without Meta-analysis [36] reporting guidelines. The review protocol was prospectively registered with PROSPERO (the International Prospective Register of Systematic Reviews).

Search Strategy

Relevant peer-reviewed papers published in English were identified by comprehensively searching key electronic databases from the earliest available evidence to October 27, 2021. These were the Web of Science, Cochrane Central Register of Controlled Trials, CINAHL, MEDLINE, Embase, PsycINFO, Scopus, ClinicalTrials.gov, and EBSCO Open Dissertations databases. The backward snowballing technique was also used to identify potentially relevant papers [37]. This involved looking at the reference lists of the selected papers to see if any more papers could be identified. The search strategy was created and refined based on consultation with a library information specialist (in Multimedia Appendix 1).

Inclusion and Exclusion Criteria

The inclusion and exclusion criteria used to screen the articles were based on PICOS (population, intervention, comparator, outcome, and study design) [38].

Population

Studies were only included if the pregnant women were at least 18 years old. Studies were excluded if they combined samples of pregnant and nonpregnant women, male partners, or those aged younger than 18 years. Given that we focused on the broad spectrum of interventions for depression (prevention [universal, selective, indicated], treatment [case identification, standard treatment for known disorders], and maintenance [compliance with long-term treatment, aftercare] [22]), this may include a population of pregnant women with different depression risk groups.

Intervention

We focused on digitalized CBT interventions that primarily used technological platforms to target depression, such as the internet, computers, mobile phones, and virtual reality. The term digitalized CBT was based on 2 previously established definitions. First, we used the umbrella term digital, as it covers the full spectrum of digital technology, such as the internet, electronic devices, and mobile phones [39]. Second, we referred to the definition of technology-empowered CBT outlined by Wolters et al [40]: "CBT-based interventions integrating technology varying from basic online bibliotherapy to online self-help therapy, therapist-supported computerized CBT, smartphone applications (apps), traditional CBT delivered via telephone or videoconferencing, and combinations of these forms." We only included interventions that began during pregnancy, but some of them also provided follow-up support during the postnatal period. We included any depression symptom intervention across the Institute of Medicine's spectrum of interventions for mental health based on the risk of the target population: prevention (universal, selective, indicated), treatment (case identification, standard treatment for known disorders), and maintenance (compliance with long-term treatment, aftercare) [22]. Both guided and nonguided interventions were included. We excluded studies on psychosocial interventions that began postnatally and those that primarily evaluated nondigitalized CBT interventions, such as face-to-face group interventions and those with telephone support but no digital element.

Comparisons

Studies that compared digitalized interventions with standard treatment, waiting lists for treatment, and secondary or other interventions were included.

Outcome

Studies were included if the primary outcome was at least one validated measure of symptoms of depression during the perinatal period.

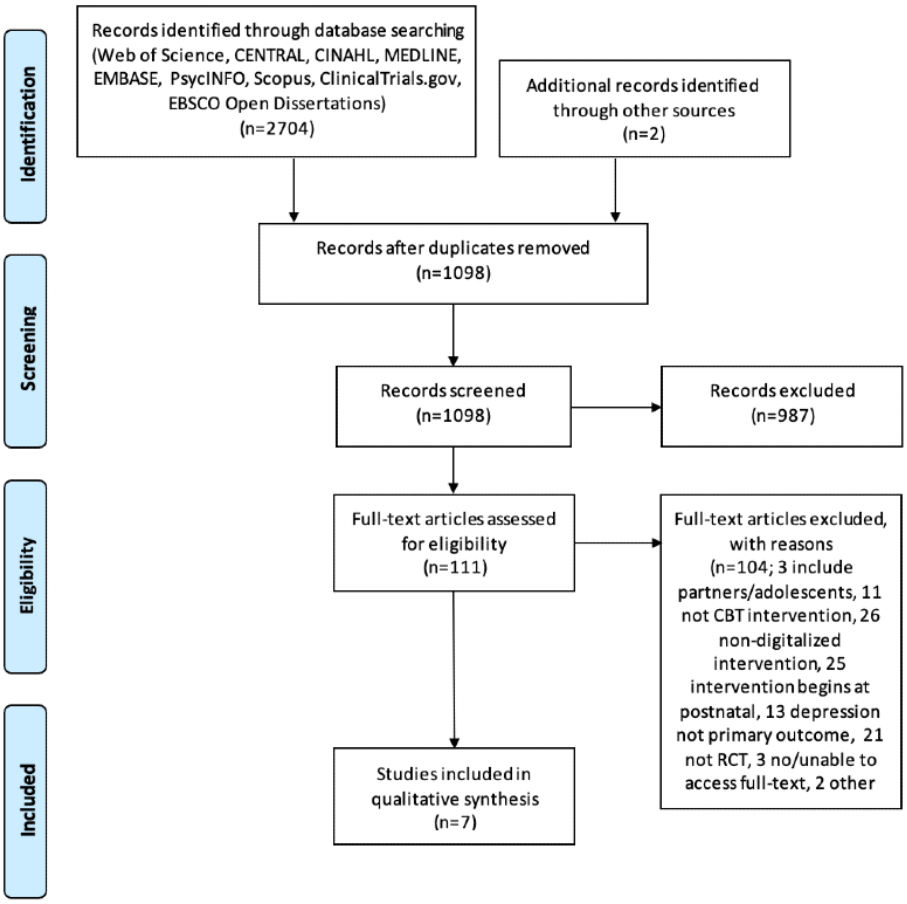
Study Design

Our review only included RCTs, as this research design is considered to be the gold standard for studying the effectiveness of interventions. RCTs are also considered to be the cornerstone for evidence-based practice [41], as they can be used to inform policy and practice decisions [42]. Only studies with full texts published in English were included. Nonrandomized, quasi-experimental, or pure qualitative studies were excluded and so were research protocols and case studies.

Study Selection and Retrieval Process

Figure 1 displays the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram for the screening and study selection processes. The search and screening processes were independently conducted by 2 reviewers (WMAWMY and OW) based on the titles and abstracts and after any duplicates were removed. Any disagreements were discussed with a senior researcher (HMM). Next, the 2 reviewers independently conducted full-text assessments based on the predefined inclusion and exclusion criteria. Both reviewers cross-checked the papers that were suggested for inclusion, and any disagreements were discussed and resolved with the senior researcher (HMM) and professor (AS). The list of excluded studies with their reasons for exclusion are available in Multimedia Appendix 2.

Figure 1. PRISMA (Preferred Reporting Items of Systematic Reviews and Meta-analyses) flow diagram.



Quality and Risk of Bias Assessment

The methodological quality of the included studies was assessed using the Revised Cochrane risk-of-bias tool for randomized trials (ROB2) [43]. This tool addresses specific domains that can influence the risk of bias in an RCT. It asks a series of questions to enable users to evaluate the potential risk of bias in RCTs based on 5 domains: the randomization process, deviations from the intended intervention, missing outcome data, outcome measurements, and how the reported results were selected. The process evaluates each domain and enables the user to reach an overall judgement, namely a low or high risk of bias or some concerns. According to the tool guidelines, identifying a high risk of bias for any individual domain will produce an overall high risk of bias. Some concerns about any individual domain will lead to the paper being categorized as some concerns or a high overall risk [43]. The risk of bias assessments were independently conducted by 2 reviewers (WMAWMY and OW). Any discrepancies were discussed and agreed with the senior researcher (HMM) and professor (AS). The specific reasons for the 3 categorizations, namely low, some concerns, or high, were recorded in a Microsoft Excel spreadsheet template provided with the ROB2 tool.

Data Extraction and Synthesis

A customized data extraction Excel spreadsheet was used by the reviewers. Data that were extracted from each of the final 7 papers included the author, year of publication, country, recruitment, sample size, and randomization. We also extracted the name, type, and mode of the intervention, screening method, primary depression scale, other mother-related outcomes, infant-related outcomes, number of sessions and duration, intervention start and end points, funding sources, brief synopsis, CBT elements, type of therapeutic guidance, and other support, such as technical or peer support. The research team also extracted and calculated the between-group effect sizes and 95% CIs after the intervention, user feedback, and the postintervention attrition rates. The Hedges g effect sizes were either extracted, if provided, or calculated using the 2020 Effect Size Calculator from the Center for Evaluation and Monitoring [44]. We extracted the means, standard deviations, and sample sizes of the intervention and control groups at the postintervention assessment point, during pregnancy. The effect sizes were interpreted as small (≤0.32), moderate (0.33 to 0.55), or large (≥0.56) [45]. The acceptability of the interventions was assessed according to the Theoretical Framework of Acceptability [46]. This suggests that intervention acceptability can be assessed at 3 time points: before, during, and after an intervention has been delivered [46]. In addition, the attrition rate percentage was calculated using the total number of subjects dropped out in the intervention group post-intervention over the total number of subjects in the intervention groups at baseline, after randomization. The data extraction was conducted by the first author (WMAWMY) and then independently checked by the second reviewer (OW) for accuracy and completeness. Disagreements were resolved through consensus discussion with a senior researcher (HMM). WMAWMY also

contacted authors of included studies with a request for missing or additional information (in Multimedia Appendix 3).

Results

The electronic searches and hand searches yielded 2706 titles, and 7 were included in this study after the duplicates were removed and the screening process was carried out [47-53]. All of the papers were published between 2015 and 2021, and 5 were published in 2019 or 2021. Figure 1 provides a flow diagram of the screening processes and study selection.

Characteristics of the Included Studies

Table 1 summarizes the characteristics of the 7 studies: 2 were from the United States, 2 were from China, and there was 1 each from Australia, Norway, and Sweden. The papers covered a total of 2830 participants, and the sample sizes ranged from 25 to 1342 at the time of randomization. The 7 studies assessed 7 different interventions. The Swedish Internet Cognitive Behavior Therapy for Antenatal Depression program had 42 participants, and the intervention, which was delivered via a secure online platform, consisted of reading material, assessments, homework, and worksheets [49]. Videos, graphics, and an online diary book were used by the Chinese Internet-based Mindful Self-Compassion Program with 314 participants [50]. The Mothers and Babies Internet Course had 852 participants, making it the second largest study. Coordinated from the United States, but with a worldwide reach, it provided fully automated online lessons that included information pages, short audio and video clips, images, and worksheets [47]. The American-based Sunnyside Group-based Internet Intervention was an online website with various interactive tools and games that facilitated learning and peer interaction. It was the smallest study, with 25 participants [48]. Mamma Mia was by far the biggest program, with 1342 Norwegian participants. It was a fully web-based intervention that combined text, pictures, pre-recorded audio files, and user input [51]. The Australian MUMentum Pregnancy Program had 87 participants. This featured a virtual online clinic and used an illustrated story centered around fictional characters [52]. The final study was the Chinese smartphone-based Mindfulness Behavioral Cognitive Therapy, with 168 participants, delivered through a smartphone app. It consisted of thematic curriculum provided through text, audio, and visual materials [53].

Two of the studies, from the United States [47] and Norway [51], employed a universal approach, and the other 5 studies used an indicated approach. Three studies recruited participants directly from hospitals: the Chinese study from a reproductive mental health multidisciplinary clinic [50], another Chinese study from the obstetrics clinic of a tertiary hospital [53], and the largest Norwegian study via midwives during regular maternal check-ups [51]. The other 4 studies combined diverse recruitment strategies, using advertisements on social media and online forums and in newspapers, mass emails, Google Ads, and flyers distributed at maternity clinics and hospitals. None of the studies screened participants using population-based approaches. Four studies screened participants during recruitment using different measures and cut-off scores. The Swedish study used a Montgomery-Åsberg Depression Rating Scale (MADRS) [54] score of 5 to 35 and no or a low risk of suicide [49]. The Chinese study used an Edinburgh Postnatal Depression Scale (EPDS) [55] score of ≥9 [50], while the other Chinese study used an EPDS score of ≥9 or Patient Health Questionnaire (PHQ)-9 [56] of ≥4 [53]. The small American study used PHQ-8 scores of 5 to 14, which indicated no diagnosis of major depression [48]. When it came to measuring depression, 6 studies used the EPDS, as either the only measure or one of the outcome measures. However, the EPDS outcome cut-off scores used for probable depression differed: 4 studies used >9 points [47,50,51,53], and 2 studies used >12 points [49,52]. Other studies used the PHQ-9, Center for Epidemiological Studies-Depression [57], MADRS-Self Report, and Beck Depression Inventory II [58] to measure depression outcomes. The studies also assessed various other mother-related outcomes, but only 2 studies reported infant-related or mother-infant attachment outcomes. These were the Chinese study [50], which used the Infant Behavior Questionnaire Revised-Very Short Form [59] and the Australian study [52], which used the Maternal Antenatal Attachment Scale [60]. Five studies [48-50,52,53] also assessed anxiety as secondary outcomes, using the Generalized Anxiety Disorder 7 scale [61], the State-Trait Anxiety Inventory [62], or the Inventory of Depression and Anxiety Symptoms [63].

Table 1. Summary of included studies.

Author, reference and country	Recruitment	Intervention name	Randomization and sample size	Intervention type and mode	Screening	Primary depression scale	Other mother-related outcomes	Infant-related outcomes
Guided								
Forsell et al [49], Sweden	Advertisements on social media, blogs, online forums, in newspapers, and flyers distributed in maternity clinics	Internet Cognitive Behavior Therapy for Antenatal Depression	Intervention: n=22; controls: n=20; total: n=42	Indicated, internet and website for intervention group with treatment as usual for controls	MADRS-S[a] scores of 15 to 35 and no/low risk of suicide	MADRS-S	EPDS[b] (>12), WSAS[c], GAD[d], ISI[e], EQ-5D-3L[f], AUDIT[g], DUDIT[h]	-[i]
Guo et al [50], China	Pregnant women attending Tianjin First Center Hospital	Internet-based Mindful Self-Compassion Program	Intervention: n=157; controls: n=157; total: n=314	Indicated, internet and website for intervention group with waiting-list control group	EPDS (≥9)	EPDS (>9)	STAI[j]-I and II, BDI[k]-II, Chinese Mindfulness Attention Awareness Self-Compassion Scale, WHO-5[l], PSI[m], Comprehensive Parenting Behavior Questionnaire	IBQ-R VSF[n]
Unguided								
Barrera et al [47], United States	Search engine directories and Google Ads Worldwide reach	Mothers and Babies Internet Course	Intervention: n=435; controls n=417; total: n=852	Universal, internet, and website for intervention group with information only for controls	N/A[o]	CESD[p]	EPDS (>9), Major Depressive Episode Screener	-
Duffecy et al [48], United States	Pregnant women invited after being identified from electronic records of a university hospital; mass emails and advertisements	Sunnyside Group-based Internet Intervention	Intervention: n=18; controls: n=7; total: n=25	Indicated, internet and website for intervention group with extra features for intervention group and just online group for controls	PHQ[q] scores of 5 -14 and no diagnosis of major depression	PHQ-8	HDRS[r], IDAS[s], SCID-I[t], MINI[u] Suicide	-
Haga et al [51], Norway	Women attending routine prenatal care in hospital clinics	Mamma Mia	Intervention: n=678; controls: n=664; total: n=1342	Universal, internet and website for intervention group, with usual perinatal care for controls	N/A	EPDS (>9)	-	-
Loughnan et al [52], Australia	Advertisements on social media, online forums, and flyers distributed in maternity hospitals	MUMentum Pregnancy Program	Intervention: n=43; controls: n=44; total: n=87	Indicated, internet and website for intervention group, with treatment as usual for controls	Met the criteria for a probable diagnosis of generalized anxiety disorder and/or major depressive disorder	PHQ-9	GAD-7, Kessler 10-item Psychological Distress scale, EPDS (>12), WHO-QOL[v], 9-item BDI-II	MAAS[w]
Sun et al [53], China	Obstetrics clinic of a tertiary hospital in Jinan, Shandong	Spirits Healing (in Chinese) app	Intervention: n=84; controls: n=84; total: n=168	Indicated, internet and smartphone app for intervention group, with regular WeChat health consultations for controls	EPDS (≥9) or PHQ-9 (≥4)	EPDS	GAD-7, PSS[x], PANAS[y], PSQI[z], PRMQ[aa], W-DEQ[ab]	-

[a]MADRS-S: Montgomery-Åsberg Depression Rating Scale-Self report version.
[b]EPDS: Edinburgh Postnatal Depression Scale.
[c]WSAS: Work and Social Adjustment Scale.
[d]GAD: Generalized Anxiety Disorder.
[e]ISI: The Insomnia Severity Index.
[f]EQ-5D-3L: EuroQoL 5-Dimension 3-Level.
[g]AUDIT: Alcohol Use Disorders Identification Test.
[h]DUDIT: Drug Use Disorders Identification Test.
[i]None reported.
[j]STAI: State-Trait Anxiety Inventory.
[k]BDI: Beck Depression Inventory.
[l]WHO-5: World Health Organization 5 Well-Being Index.
[m]PSI: Parenting Stress Index.
[n]IBQ-R VSF: Infant Behavior Questionnaire Revised-Very Short Form.
[o]N/A: not applicable.
[p]CESD: Center for Epidemiological Studies-Depression.
[q]PHQ: Patient Health Questionnaire.
[r]HADRS: Hamilton Depression Rating Scale.
[s]IDAS: Inventory of Depression and Anxiety Symptoms.
[t]SCID-I: Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders Axis-I Disorders.
[u]MINI: Mini International Neuropsychiatric Interview.
[v]WHO-QOL: World Health Organization Quality of Life scale.
[w]MAAS: Maternal Antenatal Attachment Scale.
[x]PSS: Perceived Stress Scale.
[y]PANAS: Positive and Negative Affect Schedule.
[z]PSQI: Pittsburgh Sleep Quality Index.
[aa]PRMQ: Prospective and Retrospective Memory Questionnaire.
[ab]W-DEQ: Wijma Delivery Expectancy Questionnaire.

Interventions Used by the Studies

Table 2 summarizes the description of each intervention, while Figure 2 compares the earliest starting point to the latest end point for each intervention. All the digitalized study interventions were delivered on the internet but varied in terms of their session structure and duration, the type of support provided, and the duration. Six interventions were delivered via a website, and only 1 [53], which was recently published, was delivered via a smartphone app. The shortest intervention, from Australia, lasted for 4 weeks and was comprised of 3 brief unguided internet-based CBT (iCBT) sessions [52], while the largest Norwegian intervention, Mamma Mia, had the longest duration, of 11.5 months, and provided sessions during pregnancy and postnatal periods [51]. Two of the 7 interventions were guided, as they incorporated an element of therapeutic guidance. The Swedish intervention was provided by CBT-trained therapists, who were not mental health professionals, on a fully digital platform [49], but no telephone guidance was provided. The level of therapeutic guidance provided by this guided iCBT varied, and messages sent via the internet platform were responded to within 48 hours during weekdays. The Chinese intervention did not detail the type of guided instructions that were provided [50]. In addition to therapeutic guidance, several of the guided and unguided interventions employed other forms of support, such as peer support discussion groups or platforms, as well as technical support by the research team members. Table 3 shows a brief synopsis and the CBT elements of the 7 interventions.

Table 2. Summary of the interventions in the included studies.

Author(s) and country	Intervention name	Age (years), mean (SD), range	Type of therapeutic guidance	Other support (eg, technical or peer)	Funder
Guided					
Forsell et al [49], Sweden	Internet Cognitive Behavior Therapy for Antenatal Depression	Intervention: 31.2 (3.7); controls: 30.8 (5.3)	Personalized feedback using written online messages. Therapists only had basic CBT[a] training and no prior experience nor any special education or training in order to treat this specific population	Anonymous online discussions between participants	Swedish Research Council, regional agreement between Karolinska Institutet and Stockholm City Council, and regional agreement between Umeå University and Västerbotten County Council (ALF)
Guo et al [50], China	Internet-based Mindful Self-Compassion Program	Intervention: 29.8 (6.2); controls: 31.4 (5.7)	Detailed information not available	No information provided	No information provided
Unguided					
Barrera et al [47], United States	The Mothers and Babies Internet	Intervention: 29.81 (6.09), 18-43; controls: 30.59 (4.99), 19-42	N/A[b]	No information provided	National Institute of Mental Health, Robert Wood Johnson Health Disparities Seed Grant, University of California Committee on Latino Research, and SFGH[c] Department of Psychiatry
Duffecy et al [48], United States	Sunnyside Group-based Internet Intervention	Intervention: 30.5 (4.05), 25-45; controls: not provided	N/A	Peer support and contact moderator tool for technical or group issues	NIMH, National Center for Advancing Translational Sciences of the NIH[d]
Haga et al [51], Norway	Mamma Mia	Intervention: 31.0 (4.6); controls: 31.1 (4.5)	N/A	Notification to talk to someone or seek professional help when the presence of some or many depression symptoms	Research Council of Norway
Loughnan et al [52], Australia	MUMentum Pregnancy Program	Intervention: 31.69 (4.44); control: 31.54 (3.63)	N/A	Technical assistance by research technicians	NHMRC[e], HCF[f] Research Foundation, Rotary Health Australia, and the David Henning Memorial Foundation
Sun et al [53], China	Spirits Healing (in Chinese) app	Intervention: 30.27 (3.80); control: (29.55) 4.21	N/A	No information provided	Chinese National Funding of Social Sciences and China Scholarship Council

[a]CBT: cognitive behavioral therapy.

[b]N/A: not applicable.

[c]SFGH: San Francisco General Hospital.

[d]NIH: National Institutes of Health.

[e]NHMRC: National Health and Medical Research Council.

[f]HCF: Hospitals Contribution Fund.

Figure 2. Comparison of interventions from the earliest starting point to the latest endpoint.

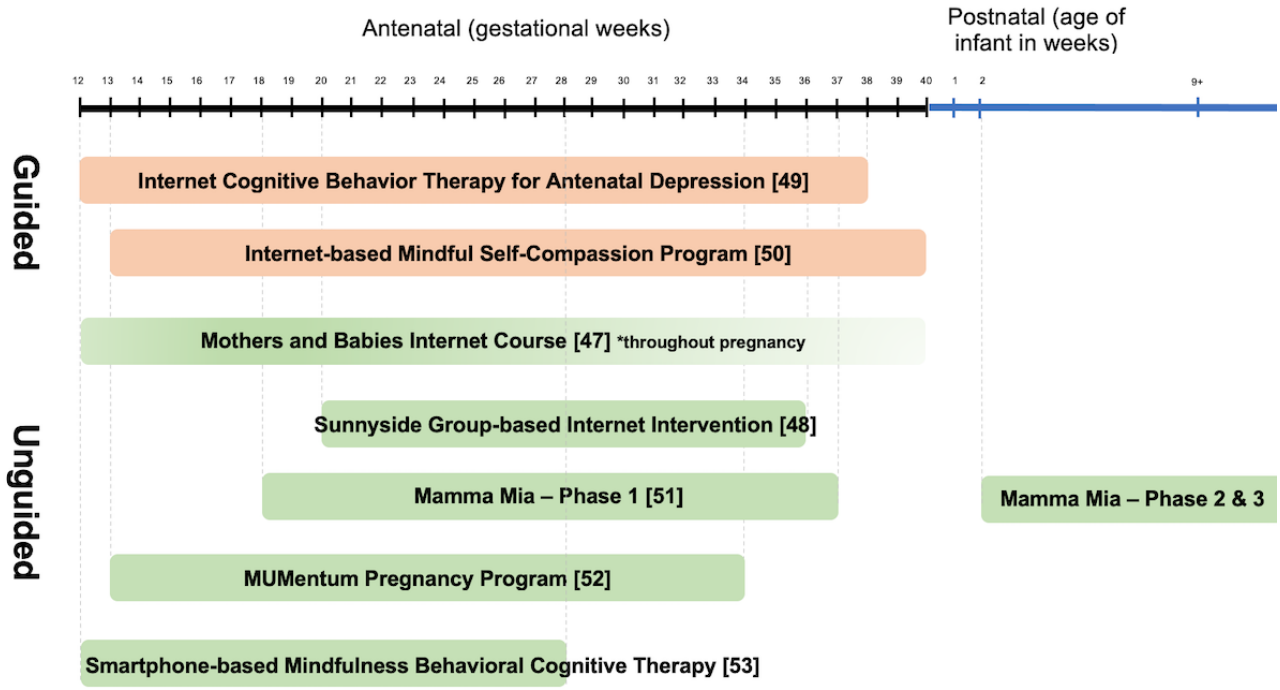


Table 3. Brief intervention synopsis, themes, and cognitive behavioral therapy (CBT) elements.

Author(s) and country	Brief synopsis	Themes	CBT elements
Guided			
Forsell et al [49], Sweden	The iCBT[a] Internet Psychiatry Clinic is an intervention for antenatal depression and is an adapted version of iCBT for depression. The platform is a form of guided self-help treatment consisting of reading material (about 75,000 words), assessments, homework, and worksheets via a secure online platform. The platform can be accessed anytime and anywhere using a computer or mobile device with an internet connection. There are 10 modules of guided self-help by nonexpert therapists trained in CBT.	Introduction, being pregnant, behavioral activation, cognitive restructuring, relationships, anxiety and worry, sleep problems, and summary and relapse prevention	Psychoeducation (depression, CBT, myths, facts, and physiological changes), behavioral activation (positive and negative reinforcement behaviors), cognitive restructuring (negative automatic thoughts, acceptance, cognitive biases), psychoeducation (relationships, communication, role transition, anxiety and fear of labor and sleep), and homework
Guo et al [50], China	Mindful Self-Compassion Program (MBSP) is aimed at promoting self-regulatory skills of pregnant women at high risk for postpartum depression focusing on self-compassion. The program utilizes videos involving different types of exercises with guided instructions sequentially provided after completion of a previous module. The 6-week program lasts 10 hours with guided instructions: 36 episodes each lasting about 15 minutes	Not available	Largely based on mindfulness CBT; understanding and applying self-compassion, skills to manage difficult emotions rather than solving specific problems, exercises with guided instructions. Users were encouraged to practice the skills during the day and provided with an online diary book for reflection.
Unguided			
Barrera et al [47], United States	The Mothers and Babies Internet (e-MB) is aimed at Spanish- and English-speaking pregnant women to reduce the risk of postpartum depression. The e-MB consists of fully automated lessons and is sequential in nature, whereby each session needed to be completed first before proceeding to the next lesson. Following completion of each session, participants may access the lesson and worksheets infinitely for review. The e-MB consisted of 8 flexible sessions of fully automated self-help.	Not available	Information pages, audio, video, images, and worksheets based on the cognitive behavior framework, social learning theory, reality management training, attachment theory, and diverse sociocultural issues
Duffecy et al [48], United States	The Sunnyside website group intervention consisted of 10–15-minute didactic lessons using text and video material and included several interactive features in the form of (1) an activity feed, which was a constantly updating feed that displayed each of the women's activity on the site, whereby other participants can like and comment or provide feedback to other women's posts, (2) discussion questions posted by study staff after each session to encourage interaction, (3) individual garden plot and community garden linked to individual's profiles and providing gamification and interactive features individually and between other users, (4) contact moderator tool to report any issues with the site or with group members. After each session, participants are prompted with a "call to action" encouraging them to apply the skills learned during the sessions. The intervention consisted of an 8-week unguided online program.	Your mood and your pregnancy, worries about me and my baby, mood management, challenging your thinking, positive activity during pregnancy, physical activity during pregnancy, partner communication and support, body image and sex during pregnancy and postpartum, relationships with your mother and mother-in-law, challenges in relationships with friends and others, monitoring kick counts and other pregnancy anxiety, anxiety and parenthood, relaxation, employment issues, during and after the birth, moving forward, and conclusions	Psychoeducation, mood management, thought challenging, positive and physical activities, relationship with partner and others in the social circle, anxiety during pregnancy and parenthood, relaxation, employment issues and managing resources during and after labor; 5 interactive CBT tools: thought restructuring (think), mood tracking (feel), activity scheduling and monitoring (do), relaxation (relax), and goal setting (achieve)

Author(s) and country	Brief synopsis	Themes	CBT elements
Haga et al [51], Norway	Mamma Mia is a free universal preventive intervention for perinatal depression. The intervention is delivered by email and interactive websites, combining text, pictures, prerecorded audio files, and user input. User receives an email with a hyperlink for each session that lasts around 10 minutes each. The hyperlink directs the user to the Mamma Mia web page, and the intervention content proceeds sequentially to the next web pages (tunnel information architecture) to ensure continuity of the program narrative. The intervention consists of 3 fully automated phases with 44 sessions over 11.5 months: phase 1 (during pregnancy); phase 2 (infant is 2-3 weeks old, for 6 weeks); and final phase (10 sessions over an 18-week period)	Knowledge, expectancies and attitudes, attachment, emotion regulation, and help-seeking, relationship satisfaction, and communication skills	Assessment of depressive symptoms, metacognitive therapy, positive psychology, couples' therapy, breastfeeding, and psychoeducation; the metacognitive element emphasized the process of inflexible and recurrent thinking style due to negative thoughts, feelings, or beliefs. Acceptance commitment therapy and mindfulness elements were also incorporated.
Loughnan et al [52], Australia	The MUMentum Pregnancy transdiagnostic intervention targets anxiety and depression symptoms and is delivered via the online Virtual Clinic system. The program emphasized a short, illustrated story centered around 2 fictional characters experiencing depression and anxiety during pregnancy. The characters learn to manage their symptoms by applying CBT skills in the context of the character experiences, challenges, and symptoms common during pregnancy. The system employs a 7-day lockout period implemented between lessons to ensure participants spend time revising and implementing the lesson material before moving onto the next lesson. Of this, participants are also notified via email and SMS reminders regarding new lessons and to stay on schedule. The 4-week program consists of 3 brief unguided self-help lessons with only technical support.	About anxiety and depression, identifying symptoms, cognitive behavioral model, prioritizing self-care, physical symptoms, partners and supporters, controlled breathing, progressive muscle relaxation, about thoughts, identifying unhelpful thoughts, shifting unhelpful thoughts, accepting uncertainty, thought challenging, coping cards, structured problem-solving, unhelpful behaviors (low activity; avoidance), facing your fears, activity planning and monitoring, graded exposure, assertive communication, relapse prevention	Transdiagnostic intervention for depression and anxiety during pregnancy; involved psychoeducation, cognitive restructuring, problem-solving, behavioral activation, and relapse prevention; each lesson illustrated the characters' stories: introduction to core CBT skills, summary, and action plan to implement the skills and several supplementary resources.
Sun et al [53], China	The Spirits Healing (in Chinese) app is a mindfulness training program for use during pregnancy for perinatal depression and other mental health problems. It was available for Android and iOS operating systems in mainland China. The app provides reading materials, recordings for guided practice, videos, and a mindfulness journal that can be accessed anytime and utilized at users' own pace. Weekly reminder messages were sent via WeChat for users to complete the training. Participants were awarded 2 yuan (US $0.30) for completion of each week of training or each completion of assessment. The 8-week mindfulness CBT training automatically updated every day, and participants practiced according to their own schedules.	Understand mindfulness, be in the present, be mindful of negative emotions, accept difficulties, thoughts are just thoughts, enjoy daily happiness, mindful pregnancy and childbirth, continued mindfulness practice	Following psychoeducation information via the app, formal mindfulness training techniques were introduced. Users are encouraged to continue to practice, and they are supplemented with recordings and a mindfulness journal. Formal training included body scan, mindful breathing, mindful stretching, and mindful meditation. Informal training included encouragement to practice every day, pausing in the midst of daily life, mindful eating, mindful walking, and 3-minute breathing practices.

[a]iCBT: internet-based CBT.

Assessment of Risk of Bias

Figure 3 and Figure 4 display the risk of bias assessment results for each study. Six studies had a high overall risk of bias, as a high risk of bias was recorded in at least one domain, while 1 study recorded some concern for overall risk of bias. Self-reported measures are commonly used to evaluate psychological outcomes, but there were at least some concerns about the risk of bias that these provided in 7 studies, based on the ROB2 guideline [43]. The missing outcome data domain was of some concern, given nearly half (3/7, 43%) of the studies did not provide sufficient information about the reasons for dropout between groups or analyses to address these missing data. In addition, the way that the reported results were selected was of high concern. These were particularly related to the fact that 5 of the study protocols were not registered or there were inconsistencies, without justification, between the registered protocols and the published papers. These inconsistencies included the absence of follow-up results in the published paper or differences between these and the intended analysis plan in the published protocol.

Figure 3. Percentage of risk of bias based on the 5 Revised Cochrane risk-of-bias tool for randomized trials (ROB2) domains.

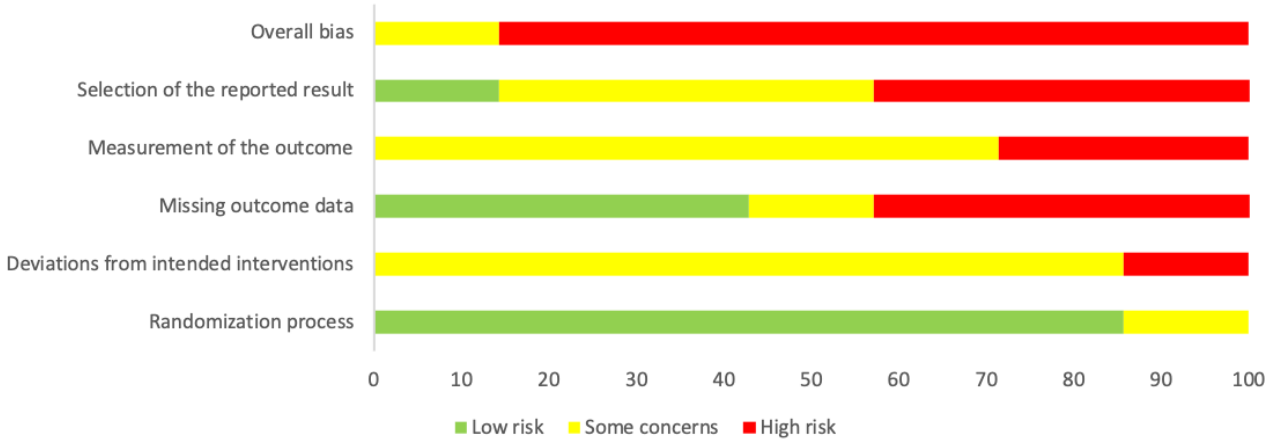
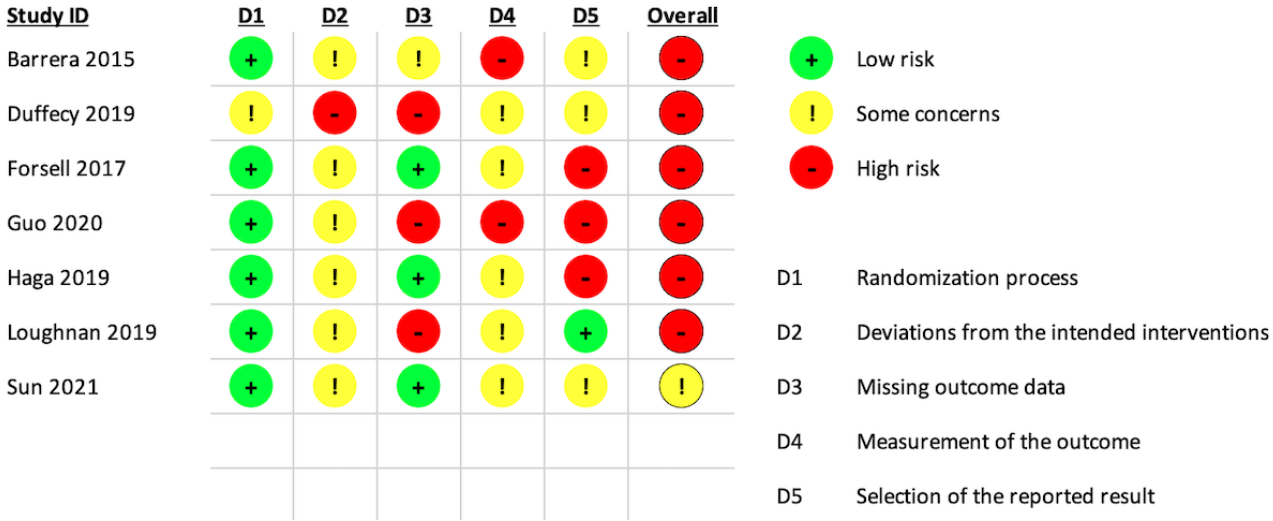


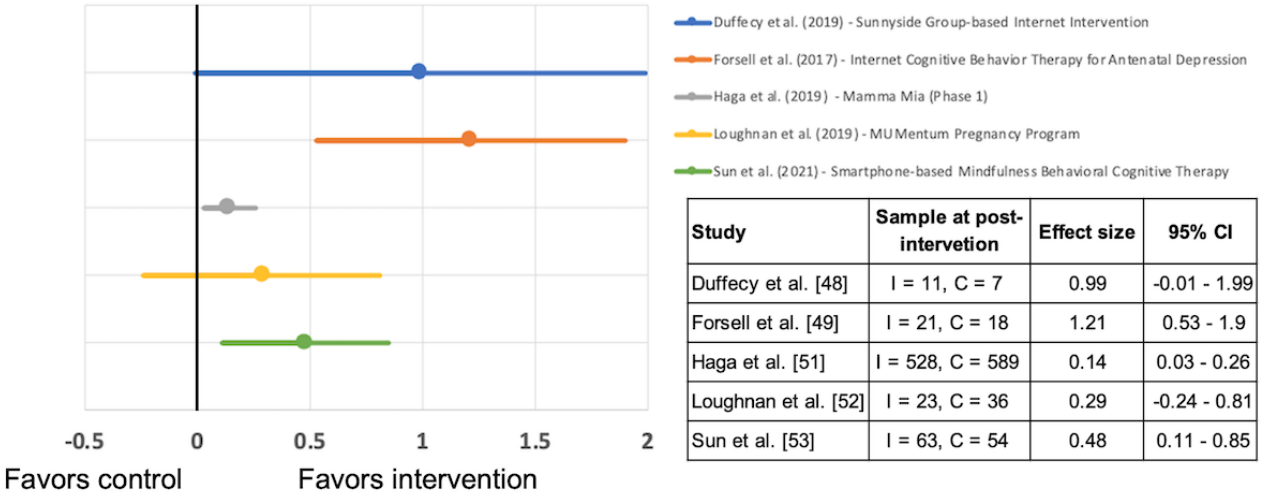
Figure 4. Risk of bias for depression outcomes, based on the individual studies.



Efficacy of Digitalized CBT Interventions for Depression Symptoms During Pregnancy

In view of the overall high risk of bias in all except one of the studies, the efficacy of the interventions needs to be interpreted with caution. Five studies provided sufficient data for effect size calculations. Figure 5 summarizes the Hedges g effect size for the 5 studies [48,49,51-53]. In general, the more positive results for the intervention than for the control groups showed promising evidence for the efficacy of digitalized CBT interventions for pregnant women. It was notable that there were differences in efficacy between the guided and unguided interventions. The highest postintervention effect size was recorded for the Swedish Internet Cognitive Behavior Therapy for Antenatal Depression program, which focused on pregnant women who were judged to have no or a low risk of suicide when screened for symptoms of depression [49]. The remaining 4 interventions were all unguided and recorded large [48], moderate [53], or small [51,52] effect sizes.

Figure 5. Hedges g effect sizes (95% CIs) for depression after programs in the intervention (I) and control (C) groups.



Study	Sample at post-intervetion	Effect size	95% CI
Duffecy et al. [48]	I = 11, C = 7	0.99	-0.01 - 1.99
Forsell et al. [49]	I = 21, C = 18	1.21	0.53 - 1.9
Haga et al. [51]	I = 528, C = 589	0.14	0.03 - 0.26
Loughnan et al. [52]	I = 23, C = 36	0.29	-0.24 - 0.81
Sun et al. [53]	I = 63, C = 54	0.48	0.11 - 0.85

Acceptability of Digitalized CBT Interventions

Table 4 compares the acceptability assessments conducted by individual studies based on 3 time points: before, during, and after the intervention was delivered. Overall, most of the participants were satisfied with the interventions they received, but high attrition rates were recorded for the 4 unguided interventions (22.1%-46.5%). The way that intervention acceptability was measured varied significantly between the studies, especially the scales that were chosen. All the studies obtained feedback after the interventions had finished. Five studies assessed acceptability during the interventions [48,49,51-53], while only 2 studies assessed acceptability before they were used [48,51]. The highest attrition rates for the unguided interventions were for the unguided MUMentum Pregnancy transdiagnostic intervention [52]. In contrast, the guided intervention recorded a very low attrition rate of 4.5% for the Swedish Internet Cognitive Behavior Therapy for Antenatal Depression program [49].

Table 4. Acceptability before, during, and after the interventions.

Author(s) and country	Acceptability before	Acceptability during	Acceptability after	Attrition
Guided				
Forsell et al [49], Sweden	N/A[a]	Treatment Credibility Scale of the Credibility/Expectancy questionnaire: good treatment credibility	Client Satisfaction Questionnaire 8: good satisfaction level; treatment adherence and utilization described	4.5%
Guo et al [50], China	N/A	N/A	Brief dropout reasons provided; attendance rate=91.8%	N/A
Unguided				
Barrera et al [47], United States	N/A	N/A	3 open-ended questions: intervention helpfulness and usefulness rated favorably; content easy to understand	N/A
Duffecy et al [48], United States	Intervention development process involved target participants; topics, site motif (visual themes and look and feel of the internet site), and usability of potential application	Use of interactive features assessed	Usability, satisfaction, and ease of use: intervention usefulness, ease of use, ease of learning, satisfaction rated favorably	38.9%
Haga et al [51], Norway	Intervention development process published in Drozd et al [64]	Dropout reasons not described in paper; other acceptability and feasibility details in paper [65]	More than half completed >80% of intervention; other acceptability and feasibility details in the paper [51]	22.1%
Loughnan et al [52], Australia	N/A	Detailed dropout reasons provided; intervention content evaluated during each session	Treatment Satisfaction Questionnaire: high satisfaction; Credibility and Expectancy Questionnaire: intervention quality rated as excellent; Intervention utilization and implementation data provided	46.5%
Sun et al [53], China	N/A	Logs of practice on formal mindfulness training	Completion rates for all 8 sessions=8.3%; completion rates for 4 sessions=52.4%	25%

[a]N/A: not available.

Discussion

Principal Findings

Our review identified 7 RCTs of digitalized CBT interventions for depression symptoms during pregnancy that were promisingly efficacious and had good acceptability. RCTs using digital CBT for depression during pregnancy have been scarce in contrast to the number of studies on digitalized CBT interventions for depression overall. We judged 6 of the 7 studies to have a high risk of bias. The high risk of bias was mainly due to missing outcome data and selection of reported results domains. These point to the need to provide accessible trial registration information with sufficient information prior to commencement of controlled trials. Although missing data in psychological intervention studies are common, a detailed description of how these missing data were addressed and reasons for dropouts in the intervention and control group need to be clarified and compared in each study. Therefore, the findings on efficacy should be considered with caution, and more high-quality studies are needed. Despite these limitations, this review produced 3 main findings. First, technology can be used to deliver CBT programs that target depression during pregnancy. Second, digitalized CBT during pregnancy showed promising evidence of efficacy and positive outcomes for depression (Hedges g for guided interventions=1.21; Hedges g for unguided interventions=0.14-0.99). Third, the digitalized CBT interventions were well-received by pregnant women, especially interventions that were guided, which had lower attrition rates than the unguided interventions (4.5% versus 22.1%-46.5%).

CBT interventions have traditionally been delivered face-to-face, but technological advances have revolutionized the way they are being adapted to various platforms. The COVID-19 pandemic has also led to an urgent increase in the need for digital care delivery. For the past 20 years, the progress in digitalized interventions has been rapid, and various evidence-based digitalized interventions have been developed and disseminated [66], especially for depression [67]. However, little is known about digitalized CBT interventions for depression during pregnancy. All 7 interventions identified in this review used the internet and website or smartphone app platforms to deliver core elements of CBT, such as psychoeducation, cognitive restructuring, problem-solving, and behavioral activation. The central themes of these were managing challenges and expectations during pregnancy by using CBT skills. These platforms contained a combination of text, images, and videos to make the interventions attractive and engaging. Other additional interactive features were unique to some interventions. For example, the American-based Sunnyside Group-based Internet Intervention included an activity feed that displayed each participant's activity on the site. This meant that other participants could like, comment, or provide feedback on other women's public posts. It also had an individual garden plot and community garden that were linked to individual profiles. The intervention provided individual and group games, where garden gnomes or flowers could be collected by completing various tasks [48]. Likewise, the Swedish Internet Cognitive Behavior Therapy for Antenatal Depression program provided an online discussion group, in which participants could interact with each other anonymously [49]. These unique interactive features allowed participants to share their lived experiences during pregnancy, as well as provided a platform for emotional peer support. These are both considered to be key functions of traditional online pregnancy forums [68].

The limited evidence and high risk of bias in 6 out of the 7 papers we reviewed prevented us from coming to firm conclusions on the efficacy of digitalized CBT for depression during pregnancy. However, our review does provide promising evidence on the efficacy and positive impact of digitalized CBT for pregnant women. Previous research has reported that the prevalence of postnatal depression symptoms can be predicted by depression symptoms during pregnancy [3]. Interventions during pregnancy allow participants to acquire and apply coping skills at an early stage, potentially before their depressive symptoms worsen. It should also be noted that 5 of the included studies used standard care, an active control, or another intervention as control groups, which may partly explain the modest effect sizes in most of the studies. The Swedish intervention [49] had an exceptionally high effect size. Unsurprisingly, that intervention used a guided and indicated approach. The intervention was delivered via an online internet platform comprised of about 75,000 words of reading material, assessments, homework, case stories, and worksheets. The therapists provided regular individual feedback, encouragement, and support in written messages, with replies within 48 hours on weekdays. Interestingly, the therapeutic guidance was provided by CBT-trained therapists who were not mental health professionals. Compelling evidence from systematic reviews and meta-analyses of iCBT research for depression has highlighted the importance of providing support from a therapist or human guidance for effective interventions [69-72].

Six out of 7 studies we reviewed also assessed other self-reported maternal outcomes: anxiety, psychological distress, alcohol and drug use, mindfulness attention awareness, general health, stress, self-compassion, well-being, quality of life and work and social adjustment, positive and negative affect, sleep-related problems, fatigue, prospective memory, retrospective memory, and fear of childbirth. Only one study assessed infant outcomes, using the Infant Behavior Questionnaire, and another study assessed mother-child attachment or bonding using the Maternal Antenatal Attachment Scale. This was somewhat surprising, as there is compelling evidence that prenatal depression has been associated with various infant-related issues, including temperament, bonding, emotional and behavioral problems, cognitive impairment, and psychopathology [73-79]. The fact that there was so little research on infant-related or mother-infant attachment outcomes by the studies we reviewed limits any insightful discussion and highlights the need for further evaluation.

All the studies we reviewed reported diverse acceptability outcomes for at least 1 of the 3 time points: before, during, and after the intervention was delivered. All the interventions recorded good acceptability, especially after the intervention had been completed, as the participants were satisfied with the interventions and felt that they were useful. Acceptability data

before and during intervention delivery are still lacking, and more acceptability components can be explored further. The Theoretical Framework of Acceptability outlines 7 components of acceptability that could be explored in intervention studies at these 3 time points: affective attitude, burden, perceived effectiveness, ethicality, intervention coherence, opportunity costs, and self-efficacy [46]. Including all the acceptability data in a single paper may be challenging due to word count limits or information overload. To deal with this issue, authors could provide acceptability evaluations in separate papers. For example, the Mamma Mia intervention included in our review was the focus of 3 separate papers on intervention development [64], acceptability and feasibility [65], and efficacy [51].

The attrition rates for the unguided interventions were high (22.1%-46.5%), and this finding was consistent with previous meta-analyses on digitalized CBT interventions for depression [70,72]. The terms "universal" and "indicated" were based on the Institute of Medicine's spectrum of mental health interventions based on the risk for the target population: prevention (universal, selective, indicated), treatment (case identification, standard treatment for known disorders), and maintenance (compliance with long-term treatment, aftercare) [22]. Our review identified 2 "universal" interventions (developed for pregnant women with any risk level for diagnosable clinical depression) and 5 "indicated" interventions (developed for pregnant women already at risk for diagnosable clinical depression [eg, those that exceed a cut-off screening score]). These classifications (ie, universal, indicated) have also been used in other perinatal distress, depression, and anxiety research [80].

Guided versus unguided and universal versus indicated approaches may have their own merits when it comes to impact. For instance, the study on the unguided and universal Mamma Mia intervention recruited the highest number of participants, 1342, with 678 in the intervention group and 664 receiving standard care. Although this study showed small effect sizes, partly due to its universal approach, more than one-half (345/678, 50.9%) of the total participants in the intervention group completed 80% or more of the year-long program, with the first phase during pregnancy and the second and third phases during the postnatal period. Any approach requires consideration based on the readiness of pregnant women to change their health behavior, particularly with reference to the Capability, Opportunity, Motivation, Behavior (COM-B) framework [81]. In general, pregnancy has been called a "teachable moment," which is described by Phelan [82] as naturally occurring life transitions or health events that are thought to motivate individuals to spontaneously adopt risk-reducing health behaviors. Although pregnancy is a time when women may benefit from increased motivation for behavioral change, based on the COM-B framework, such changes can be difficult to realize if their capabilities and opportunities are neglected [83,84]. As discussed, earlier digitalized CBT may provide the opportunity for pregnant women to access digitalized interventions. However, their lack of capacity, namely reduced energy, may prevent the women from continuously engaging with the program. That is why it is crucial for those who develop interventions to consider the 3 elements of capacity, opportunity, and motivation together, in order to minimize attrition.

Pregnancy is a relatively short period, and shorter interventions could attract pregnant women to engage. CBT is problem focused and based on the way that people think and behave, and technology can make its delivery much more cost-effective. This was proven by the Australian MUMentum Pregnancy Program included in our review [52]. The transdiagnostic intervention targeted both depression and anxiety, by using a story that centered around 2 fictional characters experiencing depression and anxiety during pregnancy. The characters learned to use CBT skills to manage their symptoms, based on their experiences and challenges during pregnancy. The program received highly positive feedback from participants, in terms of satisfaction, quality, and how they became immersed in the storyline. The small effect size recorded by this unguided intervention should not mask its potential benefits. It should be emphasized that, during the study period, participants in the treatment-as-usual control group continued to receive existing health care services. This might have had a confounding impact on their symptom levels, compared with the MUMentum Pregnancy Program intervention group [52]. Virtually all pregnant women have regular contact with health care professionals, regardless of their psychosocial well-being, and that may also have a positive effect on any depressive symptoms. This is consistent with findings of a recent individual network meta-analysis that the effectiveness of guided iCBT interventions were higher for individuals with moderate to severe depression, but unguided iCBT showed similar effectiveness for individuals with mild or subthreshold depression [72].

Limitations

Although this review was conducted with high scientific rigor and we carried out a full quality appraisal of the included studies, there are several limitations that need to be taken into account when interpreting the findings. We only included papers published in English, and this may have meant that some potential studies in other languages or grey literature from other databases were not included. All except one of the included studies had a high risk of bias. They also used participant-reported measures for most of the outcomes, which can be influenced by knowledge of the intervention being received. Using the ROB2, this resulted in at least some concern of a risk of bias for all the studies. As only 1 study reported an infant-related outcome, namely infant temperament, and only 1 study assessed mother-infant attachment, meaningful discussions about infant outcomes and maternal attachment are not possible. In addition, our review only focused on CBT and did not include interpersonal psychotherapy, although both psychotherapies were included in the US Preventive Services Task Force treatment guideline for perinatal depression that was issued in 2019 [11]. A recent review found that interpersonal psychotherapy interventions for psychological distress in perinatal populations were limited to those delivered face-to-face or by telephone [85]. The nonstandardized description of interventions in our review also made it difficult to compare them. This calls for more standardized descriptions of interventions, such as using the Template for Intervention

Description and Replication checklist [86]. Future studies could also explore various acceptability outcomes before, during, and after interventions are delivered. This could be done by using the 7 components of the Theoretical Framework of Acceptability: affective attitude, burden, perceived effectiveness, ethicality, intervention coherence, opportunity costs, and self-efficacy [46].

Conclusion

This systematic review suggests promising evidence for the potential efficacy, acceptability, and integration of digitalized CBT interventions that start during pregnancy. However, all except one of the studies included in this review recorded a high overall risk of bias. In future, we need high-quality studies with larger population-based samples to comprehensively analyze the efficacy of the interventions and explore the mechanisms of change. Digital interventions may have significant global implications when planning effective, nonstigmatizing, and cost-effective mental health treatment to prevent the long-term consequences of psychosocial problems in pregnancy. These could also have other wide-ranging clinical applications. Furthermore, there is an urgent need to study digital and remote interventions, as mental health services struggle to deal with the considerable increase in demand due to the COVID-19 pandemic.

Acknowledgments

This research was funded by Academy of Finland Flagship Programme (decision number 320162). The sponsors had no role in any aspect of the study or the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strings and databases.
[DOCX File , 23 KB - jmir_v24i2e33337_app1.docx]

Multimedia Appendix 2
List of excluded studies after full-text review.
[DOCX File , 123 KB - jmir_v24i2e33337_app2.docx]

Multimedia Appendix 3
Attempts to contact authors.
[DOCX File , 20 KB - jmir_v24i2e33337_app3.docx]

References

1. Yin X, Sun N, Jiang N, Xu X, Gan Y, Zhang J, et al. Prevalence and associated factors of antenatal depression: Systematic reviews and meta-analyses. Clin Psychol Rev 2021 Feb;83:101932 [FREE Full text] [doi: 10.1016/j.cpr.2020.101932] [Medline: 33176244]
2. Wilcox M, McGee B, Ionescu D, Leonte M, LaCross L, Reps J, et al. Perinatal depressive symptoms often start in the prenatal rather than postpartum period: results from a longitudinal study. Arch Womens Ment Health 2021 Feb;24(1):119-131 [FREE Full text] [doi: 10.1007/s00737-020-01017-z] [Medline: 32016551]
3. Faisal-Cury A, Menezes P. Antenatal depression strongly predicts postnatal depression in primary health care. Braz J Psychiatry 2012 Dec;34(4):446-450 [FREE Full text] [doi: 10.1016/j.rbp.2012.01.003] [Medline: 23429816]
4. Fekadu Dadi A, Miller ER, Mwanri L. Antenatal depression and its association with adverse birth outcomes in low and middle-income countries: A systematic review and meta-analysis. PLoS One 2020 Jan 10;15(1):e0227323 [FREE Full text] [doi: 10.1371/journal.pone.0227323] [Medline: 31923245]
5. Yedid Sion M, Harlev A, Weintraub AY, Sergienko R, Sheiner E. Is antenatal depression associated with adverse obstetric and perinatal outcomes? J Matern Fetal Neonatal Med 2016 Mar 09;29(6):863-867. [doi: 10.3109/14767058.2015.1023708] [Medline: 25777791]
6. Stein A, Pearson R, Goodman S, Rapa E, Rahman A, McCallum M, et al. Effects of perinatal mental disorders on the fetus and child. The Lancet 2014 Nov;384(9956):1800-1819 [FREE Full text] [doi: 10.1016/s0140-6736(14)61277-0]
7. Hay RE, Reynolds JE, Grohs MN, Paniukov D, Giesbrecht GF, Letourneau N, et al. Amygdala-prefrontal structural connectivity mediates the relationship between prenatal depression and behavior in preschool boys. J Neurosci 2020 Sep 02;40(36):6969-6977 [FREE Full text] [doi: 10.1523/JNEUROSCI.0481-20.2020] [Medline: 32788182]
8. Posner J, Cha J, Roy AK, Peterson BS, Bansal R, Gustafsson HC, et al. Alterations in amygdala-prefrontal circuits in infants exposed to prenatal maternal depression. Transl Psychiatry 2016 Nov 01;6(11):e935 [FREE Full text] [doi: 10.1038/tp.2016.146] [Medline: 27801896]

9. Amiel Castro RT, Pinard Anderman C, Glover V, O'Connor TG, Ehlert U, Kammerer M. Associated symptoms of depression: patterns of change during pregnancy. Arch Womens Ment Health 2017 Feb 22;20(1):123-128 [FREE Full text] [doi: 10.1007/s00737-016-0685-6] [Medline: 27878386]
10. Weissman MM. Intergenerational study of depression: a convergence of findings and opportunities. Psychol Med 2020 Jan;50(1):170-172. [doi: 10.1017/S0033291719002939] [Medline: 31647040]
11. US Preventive Services Task Force, Curry SJ, Krist AH, Owens DK, Barry MJ, Caughey AB, et al. Interventions to prevent perinatal depression: US Preventive Services Task Force Recommendation Statement. JAMA 2019 Feb 12;321(6):580-587. [doi: 10.1001/jama.2019.0007] [Medline: 30747971]
12. Le H, Perry DF, Stuart EA. Randomized controlled trial of a preventive intervention for perinatal depression in high-risk Latinas. J Consult Clin Psychol 2011 Apr;79(2):135-141. [doi: 10.1037/a0022492] [Medline: 21319897]
13. Gingrich J, Malm H, Ansorge M, Brown A, Sourander A, Suri D, et al. New insights into how serotonin selective reuptake inhibitors shape the developing brain. Birth Defects Res 2017 Jul 17;109(12):924-932 [FREE Full text] [doi: 10.1002/bdr2.1085] [Medline: 28714607]
14. Malm H, Brown AS, Gissler M, Gyllenberg D, Hinkka-Yli-Salomäki S, McKeague IW, et al. Gestational exposure to selective serotonin reuptake inhibitors and offspring psychiatric disorders: a national register-based study. J Am Acad Child Adolesc Psychiatry 2016 May;55(5):359-366 [FREE Full text] [doi: 10.1016/j.jaac.2016.02.013] [Medline: 27126849]
15. Brown AS, Gyllenberg D, Malm H, McKeague IW, Hinkka-Yli-Salomäki S, Artama M, et al. Association of selective serotonin reuptake inhibitor exposure during pregnancy with speech, scholastic, and motor disorders in offspring. JAMA Psychiatry 2016 Nov 01;73(11):1163-1170. [doi: 10.1001/jamapsychiatry.2016.2594] [Medline: 27732704]
16. Da Costa D, Zelkowitz P, Nguyen TV, Deville-Stoetzel JB. Mental health help-seeking patterns and perceived barriers for care among nulliparous pregnant women. Arch Womens Ment Health 2018 Dec;21(6):757-764 [FREE Full text] [doi: 10.1007/s00737-018-0864-8] [Medline: 29846799]
17. Goodman J. Women's attitudes, preferences, and perceived barriers to treatment for perinatal depression. Birth 2009 Mar;36(1):60-69 [FREE Full text] [doi: 10.1111/j.1523-536X.2008.00296.x] [Medline: 19278385]
18. Jones A. Help seeking in the perinatal period: a review of barriers and facilitators. Soc Work Public Health 2019;34(7):596-605 [FREE Full text] [doi: 10.1080/19371918.2019.1635947] [Medline: 31242074]
19. Kopelman RC, Moel J, Mertens C, Stuart S, Arndt S, O'Hara MW. Barriers to care for antenatal depression. Psychiatr Serv 2008 Apr;59(4):429-432 [FREE Full text] [doi: 10.1176/ps.2008.59.4.429] [Medline: 18378843]
20. Singla DR, Lawson A, Kohrt BA, Jung JW, Meng Z, Ratjen C, et al. Implementation and effectiveness of nonspecialist-delivered interventions for perinatal mental health in high-income countries: a systematic review and meta-analysis. JAMA Psychiatry 2021 May 01;78(5):498-509. [doi: 10.1001/jamapsychiatry.2020.4556] [Medline: 33533904]
21. World Health Organization. Mental Health Atlas 2017. 2018 Jan 1. URL: https://www.who.int/publications/i/item/9789241514019 [accessed 2022-02-03]
22. Institute of Medicine (US) Committee on Prevention of Mental Disorders. In: Haggerty R, Mrazek P, editors. Reducing risks for mental disorders: Frontiers for preventive intervention research. Washington DC: National Academy Press; 1994.
23. San Martin Porter MA, Kisely S, Alati R. Screening for perinatal depression and predictors of underscreening: findings of the Born in Queensland study. Med J Aust 2019 Aug;211(4):190-190.e1 [FREE Full text] [doi: 10.5694/mja2.50253] [Medline: 31264220]
24. Kazdin AE, Blase SL. Rebooting psychotherapy research and practice to reduce the burden of mental illness. Perspect Psychol Sci 2011 Jan 03;6(1):21-37 [FREE Full text] [doi: 10.1177/1745691610393527] [Medline: 26162113]
25. Kazdin AE, Rabbitt SM. Novel models for delivering mental health services and reducing the burdens of mental illness. Clinical Psychological Science 2013 Jan 23;1(2):170-191. [doi: 10.1177/2167702612463566]
26. Ristkari T, Kurki M, Suominen A, Gilbert S, Sinokki A, Kinnunen M, et al. Web-based parent training intervention with telephone coaching for disruptive behavior in 4-year-old children in real-world practice: implementation study. J Med Internet Res 2019 Apr 11;21(4):e11446 [FREE Full text] [doi: 10.2196/11446] [Medline: 30973337]
27. Sourander A, McGrath PJ, Ristkari T, Cunningham C, Huttunen J, Lingley-Pottie P, et al. Internet-assisted parent training intervention for disruptive behavior in 4-year-old children: a randomized clinical trial. JAMA Psychiatry 2016 Apr 01;73(4):378-387. [doi: 10.1001/jamapsychiatry.2015.3411] [Medline: 26913614]
28. Sourander A, McGrath P, Ristkari T, Cunningham C, Huttunen J, Hinkka-Yli-Salomäki S, et al. Two-year follow-Up of internet and telephone assisted parent training for disruptive behavior at age 4. J Am Acad Child Adolesc Psychiatry 2018 Sep;57(9):658-668.e1 [FREE Full text] [doi: 10.1016/j.jaac.2018.07.001] [Medline: 30196869]
29. Gülmezoglu AM, Ammerdorffer A, Narasimhan M, Wilson AN, Vogel JP, Say L, et al. Self-care and remote care during pregnancy: a new paradigm? Health Res Policy Syst 2020 Sep 18;18(1):107 [FREE Full text] [doi: 10.1186/s12961-020-00627-4] [Medline: 32948188]
30. Li C, Sun X, Li Q, Sun Q, Wu B, Duan D. Role of psychotherapy on antenatal depression, anxiety, and maternal quality of life: A meta-analysis. Medicine (Baltimore) 2020 Jul 02;99(27):e20947 [FREE Full text] [doi: 10.1097/MD.0000000000020947] [Medline: 32629701]

31. Loughnan S, Joubert A, Grierson A, Andrews G, Newby J. Internet-delivered psychological interventions for clinical anxiety and depression in perinatal women: a systematic review and meta-analysis. Arch Womens Ment Health 2019 Dec;22(6):737-750 [FREE Full text] [doi: 10.1007/s00737-019-00961-9] [Medline: 31101993]
32. Lee E, Denison F, Hor K, Reynolds R. Web-based interventions for prevention and treatment of perinatal mood disorders: a systematic review. BMC Pregnancy Childbirth 2016 Feb 29;16:38 [FREE Full text] [doi: 10.1186/s12884-016-0831-1] [Medline: 26928898]
33. Nillni Y, Mehralizade A, Mayer L, Milanovic S. Treatment of depression, anxiety, and trauma-related disorders during the perinatal period: A systematic review. Clin Psychol Rev 2018 Dec;66:136-148 [FREE Full text] [doi: 10.1016/j.cpr.2018.06.004] [Medline: 29935979]
34. Smoktunowicz E, Barak A, Andersson G, Banos R, Berger T, Botella C, et al. Consensus statement on the problem of terminology in psychological interventions using the internet or digital components. Internet Interv 2020 Sep;21:100331 [FREE Full text] [doi: 10.1016/j.invent.2020.100331] [Medline: 32577404]
35. Moher D, Liberati A, Tetzlaff J, Altman D, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
36. Campbell M, McKenzie J, Sowden A, Katikireddi S, Brennan S, Ellis S, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. BMJ 2020 Jan 16;368:l6890 [FREE Full text] [doi: 10.1136/bmj.l6890] [Medline: 31948937]
37. Wohlin C. Guidelines for snowballing in systematic literature studies and a replication in software engineering. 2014 Presented at: 18th International Conference on Evaluation and Assessment in Software Engineering; May 13-14, 2014; London, United Kingdom URL: https://dl.acm.org/doi/10.1145/2601248.2601268 [doi: 10.1145/2601248.2601268]
38. Amir-Behghadami M, Janati A. Population, Intervention, Comparison, Outcomes and Study (PICOS) design as a framework to formulate eligibility criteria in systematic reviews. Emerg Med J 2020 Jun 05;37(6):387-387. [doi: 10.1136/emermed-2020-209567] [Medline: 32253195]
39. Luik AI, van der Zweerde T, van Straten A, Lancee J. Digital delivery of cognitive behavioral therapy for insomnia. Curr Psychiatry Rep 2019 Jun 04;21(7):50 [FREE Full text] [doi: 10.1007/s11920-019-1041-0] [Medline: 31161406]
40. Wolters LH, Op de Beek V, Weidle B, Skokauskas N. How can technology enhance cognitive behavioral therapy: the case of pediatric obsessive compulsive disorder. BMC Psychiatry 2017 Jun 23;17(1):226 [FREE Full text] [doi: 10.1186/s12888-017-1377-0] [Medline: 28645268]
41. Lilienfeld S, McKay D, Hollon S. Why randomised controlled trials of psychological treatments are still essential. Lancet Psychiatry 2018 Jul;5(7):536-538. [doi: 10.1016/S2215-0366(18)30045-2] [Medline: 29602738]
42. Bristow D, Carter L, Martin S. Using evidence to improve policy and practice: the UK What Works Centres. Contemporary Social Science 2015 Aug 11;10(2):126-137 [FREE Full text] [doi: 10.1080/21582041.2015.1061688]
43. Sterne J, Savović J, Page M, Elbers R, Blencowe N, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898 [FREE Full text] [doi: 10.1136/bmj.l4898] [Medline: 31462531]
44. Effect Size Calculator. Cambridge University Press & Assessment Center for Evaluation & Monitoring. URL: https://www.cem.org/effect-size-calculator [accessed 2021-08-21]
45. Lipsey MW, Wilson DB. The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. American Psychologist 1993;48(12):1181-1209. [doi: 10.1037/0003-066x.48.12.1181]
46. Sekhon M, Cartwright M, Francis J. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017 Jan 26;17(1):88 [FREE Full text] [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
47. Barrera A, Wickham R, Muñoz RF. Online prevention of postpartum depression for Spanish- and English-speaking pregnant women: A pilot randomized controlled trial. Internet Interv 2015 Sep 01;2(3):257-265 [FREE Full text] [doi: 10.1016/j.invent.2015.06.002] [Medline: 26273567]
48. Duffecy J, Grekin R, Hinkel H, Gallivan N, Nelson G, O'Hara M. A group-based online intervention to prevent postpartum depression (Sunnyside): feasibility randomized controlled trial. JMIR Ment Health 2019 May 28;6(5):e10778 [FREE Full text] [doi: 10.2196/10778] [Medline: 31140443]
49. Forsell E, Bendix M, Holländare F, Szymanska von Schultz B, Nasiell J, Blomdahl-Wetterholm M, et al. Internet delivered cognitive behavior therapy for antenatal depression: A randomised controlled trial. J Affect Disord 2017 Oct 15;221:56-64 [FREE Full text] [doi: 10.1016/j.jad.2017.06.013] [Medline: 28628768]
50. Guo L, Zhang J, Mu L, Ye Z. Preventing postpartum depression with mindful self-compassion intervention: a randomized control study. J Nerv Ment Dis 2020 Feb;208(2):101-107. [doi: 10.1097/NMD.0000000000001096] [Medline: 31868776]
51. Haga S, Drozd F, Lisøy C, Wentzel-Larsen T, Slinning K. Mamma Mia - A randomized controlled trial of an internet-based intervention for perinatal depression. Psychol Med 2019 Aug;49(11):1850-1858 [FREE Full text] [doi: 10.1017/S0033291718002544] [Medline: 30191779]
52. Loughnan S, Sie A, Hobbs M, Joubert A, Smith J, Haskelberg H, et al. A randomized controlled trial of 'MUMentum Pregnancy': Internet-delivered cognitive behavioral therapy program for antenatal anxiety and depression. J Affect Disord 2019 Jan 15;243:381-390 [FREE Full text] [doi: 10.1016/j.jad.2018.09.057] [Medline: 30266030]

53. Sun Y, Li Y, Wang J, Chen Q, Bazzano AN, Cao F. Effectiveness of smartphone-based mindfulness training on maternal perinatal depression: randomized controlled trial. J Med Internet Res 2021 Jan 27;23(1):e23410 [FREE Full text] [doi: 10.2196/23410] [Medline: 33502326]
54. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry 1979 Apr 01;134(4):382-389. [doi: 10.1192/bjp.134.4.382] [Medline: 444788]
55. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. Br J Psychiatry 1987 Jun 2;150(06):782-786. [doi: 10.1192/bjp.150.6.782] [Medline: 3651732]
56. Kroenke K, Spitzer RL. The PHQ-9: a new depression diagnostic and severity measure. Psychiatric Annals 2002 Sep 01;32(9):509-515. [doi: 10.3928/0048-5713-20020901-06]
57. Radloff LS. The CES-D Scale. Applied Psychological Measurement 2016 Jul 26;1(3):385-401. [doi: 10.1177/014662167700100306]
58. Beck AT, Steer RA, Brown GK. Manual for the Beck Depression Inventory-II. San Antonio, TX: Psychological Corporation; 1996.
59. Putnam SP, Helbig AL, Gartstein MA, Rothbart MK, Leerkes E. Development and assessment of short and very short forms of the infant behavior questionnaire-revised. J Pers Assess 2014 Nov 09;96(4):445-458. [doi: 10.1080/00223891.2013.841171] [Medline: 24206185]
60. Condon J. The assessment of antenatal emotional attachment: development of a questionnaire instrument. Br J Med Psychol 1993 Jun;66 (Pt 2):167-183 [FREE Full text] [doi: 10.1111/j.2044-8341.1993.tb01739.x] [Medline: 8353110]
61. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
62. Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA. Manual for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologists Press; 1983.
63. Watson D, O'Hara MW, Simms LJ, Kotov R, Chmielewski M, McDade-Montez EA, et al. Development and validation of the Inventory of Depression and Anxiety Symptoms (IDAS). Psychol Assess 2007 Sep;19(3):253-268. [doi: 10.1037/1040-3590.19.3.253] [Medline: 17845118]
64. Drozd F, Haga S, Brendryen H, Slinning K. An internet-based intervention (Mamma Mia) for postpartum depression: mapping the development from theory to practice. JMIR Res Protoc 2015 Oct 12;4(4):e120 [FREE Full text] [doi: 10.2196/resprot.4858] [Medline: 26476481]
65. Haga S, Drozd F, Brendryen H, Slinning K. Mamma mia: a feasibility study of a web-based intervention to reduce the risk of postpartum depression and enhance subjective well-being. JMIR Res Protoc 2013 Aug 12;2(2):e29 [FREE Full text] [doi: 10.2196/resprot.2659] [Medline: 23939459]
66. Andersson G. Internet interventions: Past, present and future. Internet Interv 2018 Jun;12:181-188 [FREE Full text] [doi: 10.1016/j.invent.2018.03.008] [Medline: 30135782]
67. Webb C, Rosso I, Rauch S. Internet-based cognitive-behavioral therapy for depression: current progress and future directions. Harv Rev Psychiatry 2017;25(3):114-122 [FREE Full text] [doi: 10.1097/HRP.0000000000000139] [Medline: 28475503]
68. Ellis L, Roberts L. Exploring the use and quality of Internet discussion forums in pregnancy: A qualitative analysis. Birth 2020 Mar;47(1):153-161 [FREE Full text] [doi: 10.1111/birt.12459] [Medline: 31583769]
69. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342 [FREE Full text] [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
70. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009;38(4):196-205 [FREE Full text] [doi: 10.1080/16506070903318960] [Medline: 20183695]
71. Spek V, Cuijpers P, Nyklícek I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. Psychol Med 2007 Mar;37(3):319-328. [doi: 10.1017/S0033291706008944] [Medline: 17112400]
72. Karyotaki E, Efthimiou O, Miguel C, Bermpohl FMG, Furukawa T, Cuijpers P, Individual Patient Data Meta-Analyses for Depression (IPDMA-DE) Collaboration. Internet-based cognitive behavioral therapy for depression: a systematic review and individual patient data network meta-analysis. JAMA Psychiatry 2021 Apr 01;78(4):361-371 [FREE Full text] [doi: 10.1001/jamapsychiatry.2020.4364] [Medline: 33471111]
73. Savory K, Garay S, Sumption L, Kelleher J, Daughters K, Janssen A, et al. Prenatal symptoms of anxiety and depression associated with sex differences in both maternal perceptions of one year old infant temperament and researcher observed infant characteristics. J Affect Disord 2020 Mar 01;264:383-392 [FREE Full text] [doi: 10.1016/j.jad.2019.11.057] [Medline: 32056774]
74. Lahti M, Savolainen K, Tuovinen S, Pesonen A, Lahti J, Heinonen K, et al. Maternal Depressive Symptoms During and After Pregnancy and Psychiatric Problems in Children. J Am Acad Child Adolesc Psychiatry 2017 Jan;56(1):30-39.e7 [FREE Full text] [doi: 10.1016/j.jaac.2016.10.007] [Medline: 27993226]
75. Räikkönen K, Pesonen A, O'Reilly J, Tuovinen S, Lahti M, Kajantie E, et al. Maternal depressive symptoms during pregnancy, placental expression of genes regulating glucocorticoid and serotonin function and infant regulatory behaviors. Psychol Med 2015 Nov;45(15):3217-3226. [doi: 10.1017/S003329171500121X] [Medline: 26096463]

76. McGrath J, Records K, Rice M. Maternal depression and infant temperament characteristics. Infant Behav Dev 2008 Jan;31(1):71-80 [FREE Full text] [doi: 10.1016/j.infbeh.2007.07.001] [Medline: 17714790]
77. Talge N, Neal C, Glover V, Early Stress, Translational Research and Prevention Science Network: Fetal and Neonatal Experience on Child and Adolescent Mental Health. Antenatal maternal stress and long-term effects on child neurodevelopment: how and why? J Child Psychol Psychiatry 2007;48(3-4):245-261 [FREE Full text] [doi: 10.1111/j.1469-7610.2006.01714.x] [Medline: 17355398]
78. Davis E, Glynn L, Schetter C, Hobel C, Chicz-Demet A, Sandman C. Prenatal exposure to maternal depression and cortisol influences infant temperament. J Am Acad Child Adolesc Psychiatry 2007 Jun;46(6):737-746 [FREE Full text] [doi: 10.1097/chi.0b013e318047b775] [Medline: 17513986]
79. Van den Bergh BRH, Mulder E, Mennes M, Glover V. Antenatal maternal anxiety and stress and the neurobehavioural development of the fetus and child: links and possible mechanisms. A review. Neurosci Biobehav Rev 2005 Apr;29(2):237-258 [FREE Full text] [doi: 10.1016/j.neubiorev.2004.10.007] [Medline: 15811496]
80. Austin M. Antenatal screening and early intervention for "perinatal" distress, depression and anxiety: where to from here? Arch Womens Ment Health 2004 Feb 1;7(1):1-6. [doi: 10.1007/s00737-003-0034-4] [Medline: 14963727]
81. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
82. Phelan S. Pregnancy: a "teachable moment" for weight control and obesity prevention. Am J Obstet Gynecol 2010 Feb;202(2):135.e1-135.e8 [FREE Full text] [doi: 10.1016/j.ajog.2009.06.008] [Medline: 19683692]
83. Olander E, Darwin Z, Atkinson L, Smith D, Gardner B. Beyond the 'teachable moment' - A conceptual analysis of women's perinatal behaviour change. Women Birth 2016 Jun;29(3):e67-e71 [FREE Full text] [doi: 10.1016/j.wombi.2015.11.005] [Medline: 26626592]
84. Olander EK, Smith DM, Darwin Z. Health behaviour and pregnancy: a time for change. J Reprod Infant Psychol 2018 Feb;36(1):1-3 [FREE Full text] [doi: 10.1080/02646838.2018.1408965] [Medline: 29517295]
85. Bright KS, Charrois EM, Mughal MK, Wajid A, McNeil D, Stuart S, et al. Interpersonal psychotherapy to reduce psychological distress in perinatal women: a systematic review. Int J Environ Res Public Health 2020 Nov 13;17(22):8421 [FREE Full text] [doi: 10.3390/ijerph17228421] [Medline: 33203014]
86. Hoffmann T, Glasziou P, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014 Mar 07;348:g1687 [FREE Full text] [doi: 10.1136/bmj.g1687] [Medline: 24609605]

Abbreviations

CBT: cognitive behavioral therapy
COM-B: Capability, Opportunity, Motivation, Behavior
EPDS: Edinburgh Postnatal Depression Scale
iCBT: internet-based CBT
MADRS: Montgomery-Åsberg Depression Rating Scale
PHQ: Patient Health Questionnaire
PICOS: population, intervention, comparator, outcome, and study design
PRISMA: Preferred Reporting Items of Systematic Reviews and Meta-analyses
RCT: randomized controlled trial
ROB2: Revised Cochrane risk-of-bias tool for randomized trials

Edited by C Basch; submitted 03.09.21; peer-reviewed by PF Zuccolo, K Matthias; comments to author 23.10.21; revised version received 12.11.21; accepted 13.12.21; published 23.02.22.

Please cite as:
Wan Mohd Yunus WMA, Matinolli HM, Waris O, Upadhyaya S, Vuori M, Korpilahti-Leino T, Ristkari T, Koffert T, Sourander A Digitalized Cognitive Behavioral Interventions for Depressive Symptoms During Pregnancy: Systematic Review J Med Internet Res 2022;24(2):e33337
URL: https://www.jmir.org/2022/2/e33337
doi:10.2196/33337
PMID:35195532

Review

Computerized Clinical Decision Support Systems for the Early Detection of Sepsis Among Adult Inpatients: Scoping Review

Khalia Ackermann[1], MPH; Jannah Baker[1], PhD; Malcolm Green[2], MN; Mary Fullick[2], BHealthSc; Hilal Varinli[3], PhD; Johanna Westbrook[1], PhD; Ling Li[1], PhD

[1]Centre for Health Systems and Safety Research, Australian Institute of Health Innovation, Macquarie University, Australia

[2]Clinical Excellence Commission, Sydney, Australia

[3]eHealth New South Wales, Sydney, Australia

Corresponding Author:
Khalia Ackermann, MPH
Centre for Health Systems and Safety Research
Australian Institute of Health Innovation
Level 6, Talavera Road
Macquarie University, 2109
Australia
Phone: 61 2 9850 2432
Email: khalia.ackermann@mq.edu.au

Abstract

Background: Sepsis is a significant cause of morbidity and mortality worldwide. Early detection of sepsis followed promptly by treatment initiation improves patient outcomes and saves lives. Hospitals are increasingly using computerized clinical decision support (CCDS) systems for the rapid identification of adult patients with sepsis.

Objective: This scoping review aims to systematically describe studies reporting on the use and evaluation of CCDS systems for the early detection of adult inpatients with sepsis.

Methods: The protocol for this scoping review was previously published. A total of 10 electronic databases (MEDLINE, Embase, CINAHL, the Cochrane database, LILACS [Latin American and Caribbean Health Sciences Literature], Scopus, Web of Science, OpenGrey, ClinicalTrials.gov, and PQDT [ProQuest Dissertations and Theses]) were comprehensively searched using terms for sepsis, CCDS, and detection to identify relevant studies. Title, abstract, and full-text screening were performed by 2 independent reviewers using predefined eligibility criteria. Data charting was performed by 1 reviewer with a second reviewer checking a random sample of studies. Any disagreements were discussed with input from a third reviewer. In this review, we present the results for adult inpatients, including studies that do not specify patient age.

Results: A search of the electronic databases retrieved 12,139 studies following duplicate removal. We identified 124 studies for inclusion after title, abstract, full-text screening, and hand searching were complete. Nearly all studies (121/124, 97.6%) were published after 2009. Half of the studies were journal articles (65/124, 52.4%), and the remainder were conference abstracts (54/124, 43.5%) and theses (5/124, 4%). Most studies used a single cohort (54/124, 43.5%) or before-after (42/124, 33.9%) approach. Across all 124 included studies, patient outcomes were the most frequently reported outcomes (107/124, 86.3%), followed by sepsis treatment and management (75/124, 60.5%), CCDS usability (14/124, 11.3%), and cost outcomes (9/124, 7.3%). For sepsis identification, the systemic inflammatory response syndrome criteria were the most commonly used, alone (50/124, 40.3%), combined with organ dysfunction (28/124, 22.6%), or combined with other criteria (23/124, 18.5%). Over half of the CCDS systems (68/124, 54.8%) were implemented alongside other sepsis-related interventions.

Conclusions: The current body of literature investigating the implementation of CCDS systems for the early detection of adult inpatients with sepsis is extremely diverse. There is substantial variability in study design, CCDS criteria and characteristics, and outcomes measured across the identified literature. Future research on CCDS system usability, cost, and impact on sepsis morbidity is needed.

International Registered Report Identifier (IRRID): RR2-10.2196/24899

(J Med Internet Res 2022;24(2):e31083) doi:10.2196/31083

KEYWORDS

sepsis; early detection of disease; clinical decision support systems; patient safety; electronic health records; sepsis care pathway

Introduction

Sepsis and Early Detection

Sepsis, defined in 2016 as "life-threatening organ dysfunction caused by a dysregulated host response to infection," is a leading cause of death worldwide [1]. A recent study by Rudd et al [2] estimated that 48.9 million cases of sepsis were reported in 2017, with 11 million sepsis-related deaths, representing 1 in 5 of all deaths globally [2]. Furthermore, survivors of sepsis often have a decreased quality of life, including higher rates of mortality, physical disabilities, chronic illnesses, mental health issues, and cognitive impairments [3-9].

Prompt administration of sepsis therapies, such as intravenous antimicrobials and fluid resuscitation, is associated with better patient outcomes and lower health care–related costs [10,11]. Therefore, it is critical to detect sepsis as early as possible to ensure rapid initiation of treatment [12-14]. Unfortunately, sepsis has no diagnostic gold standard and extremely heterogenous signs and symptoms, making it difficult for clinicians to distinguish it from other acute conditions [15]. The use of sepsis identification tools, such as the Quick Sepsis-Related Organ Failure Assessment, the National Early Warning Score, and the Adult Sepsis Pathway, helps facilitate early sepsis recognition [16-18]. However, these tools typically rely on manual input of vital sign information and score calculation by clinicians. Thus, timely sepsis identification hinges on vigilant and regular patient monitoring [19]. These difficulties often result in delayed sepsis diagnosis and treatment in hospitals [19,20].

Computerized Clinical Decision Support Systems

The extensive implementation of data-rich electronic health records in health institutions has brought the opportunity for widespread integration of digital health care support systems [21]. In particular, the incorporation of computerized clinical decision support (CCDS) into hospital systems has the potential to assist accurate and timely early sepsis detection. CCDS systems can be designed with integrated sepsis-risk warning tools that alert clinicians to patients at risk of sepsis [13,22], reducing the physical and mental workload associated with manual patient monitoring [21].

Over the past 10 years, CCDS technology has rapidly expanded, with two distinct approaches emerging: knowledge-based CCDS systems programmed with predefined rules derived from established clinical knowledge and adaptive CCDS systems using artificial intelligence and machine learning techniques [21,23,24]. In this scoping review, we focused on the use of knowledge-based CCDS systems in sepsis detection.

Research Questions and Aims

The use and implementation of sepsis CCDS systems in real-world clinical settings is a novel, rapidly expanding, and highly complex field [21,25]. In this scoping review, we systematically mapped the literature available on sepsis CCDS systems with the intention of identifying knowledge gaps and informing future research. The research question directing this review is *What is the evidence base for the use of knowledge-based clinical decision support systems in hospitals for early sepsis detection and how have they been evaluated?*

More specifically, through this scoping review, we aim to (1) scope the study contexts, designs, and research methods used; (2) summarize the study outcomes investigated; and (3) map the range of CCDS system designs and implementation features, such as sepsis clinical criteria.

Methods

Overview

The detailed methodology for conducting this scoping review was published previously in a protocol [26]. In brief, the review was guided by the Joanna Briggs Institute Reviewer's Manual [27], the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) [28], and the 5-stage scoping review framework proposed by Arksey and O'Malley [29]. A search for current reviews and protocols on this topic was undertaken and confirmed the absence of scoping reviews. A completed PRISMA-ScR checklist is attached in Multimedia Appendix 1 [28].

Study Selection

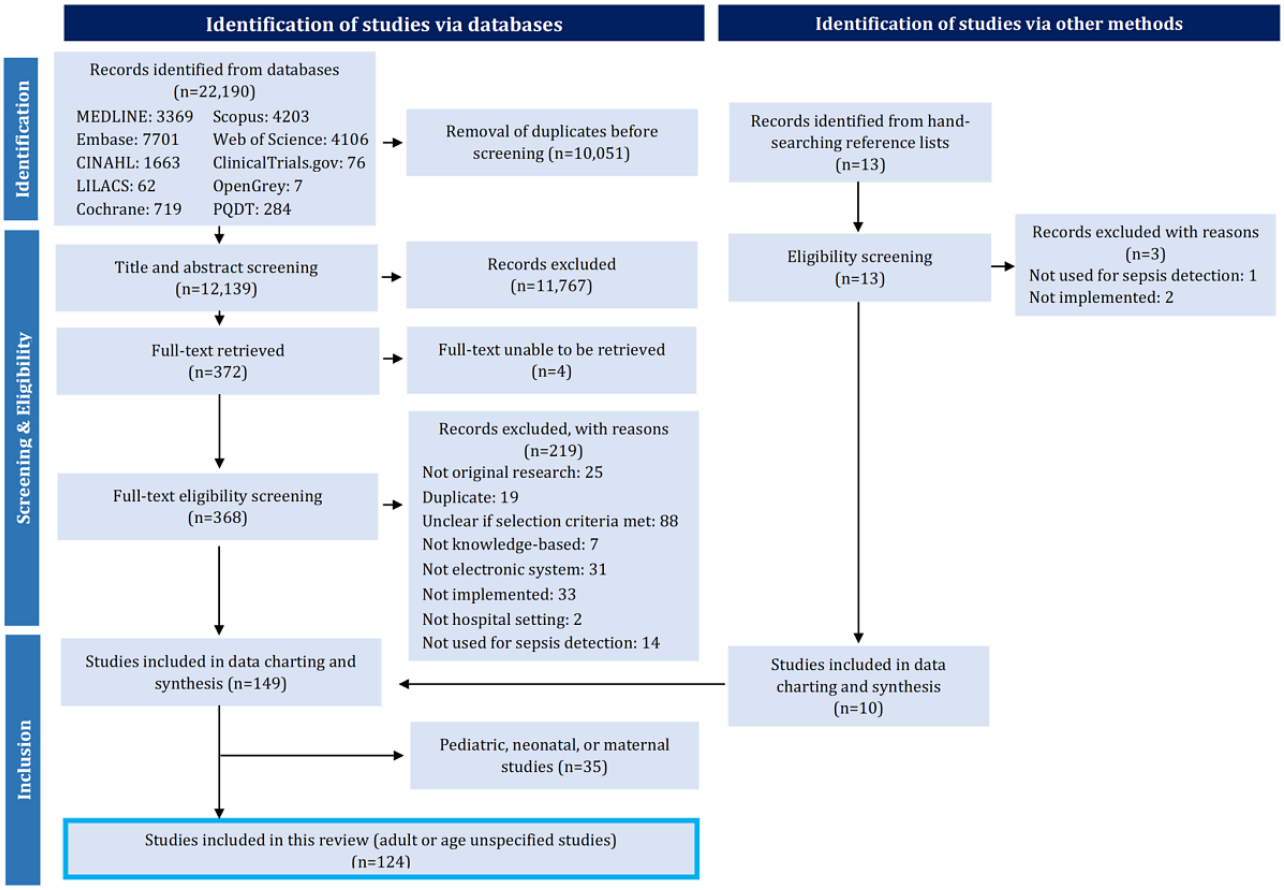
We used a broad 3-step search strategy, as outlined in our protocol [26]. An experienced librarian was consulted to help construct and refine the search. The final search strategy combined terms relating to sepsis with CCDS and detection, while excluding artificial intelligence, and was used to search MEDLINE, Embase, CINAHL, the Cochrane database, LILACS (Latin American and Caribbean Health Sciences Literature), Scopus, Web of Science, OpenGrey, ClinicalTrials.gov, and PQDT (ProQuest Dissertations and Theses Global). We restricted the search to human studies in the English language. An example of the final strategy adapted for MEDLINE can be seen in Multimedia Appendix 2. The database search was undertaken in September 2020, with no date limits applied. The reference lists of relevant systematic reviews were hand-searched to identify additional studies. Any studies identified via hand searching up until the end of data extraction (early 2021) were included. We included both peer-reviewed journal articles and gray literature (ie, conference abstracts and theses).

Following the search, duplicates were removed as was gray literature that had been published as a peer-reviewed journal article. However, we kept studies if they reported the same methods and study cohort but examined different outcomes. Using the eligibility criteria as reported in our protocol [26], 2 reviewers (KA and JB) independently performed title, abstract, and full-text screening, with any disagreements resolved through discussion or review by a third researcher (LL). Title and abstract screening was piloted with a random selection of 25 studies by both reviewers (KA and JB). Similarly, full-text

screening was piloted with a random selection of 10 studies. The 2 reviewers (KA and JB) had 100% agreement during the title and abstract screen pilot, 97.6% agreement for the full title and abstract screen, 60% agreement for the full-text screen pilot, and 77.4% agreement for the full-text screen. Hand searching was completed by 1 reviewer (KA) with identified studies confirmed by a second (JB). A PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram visually illustrating this process is shown in Figure 1.

Following screening, it was determined that the results of this review would be split over 2 papers, one investigating adult or unspecified populations and another investigating pediatric, neonatal, and maternal populations. Pediatric, neonatal, and maternal populations have remarkably different sepsis presentations and physiology compared with the general adult population [30-32]. The separation of results will allow for a more meaningful analysis. Included studies with unspecified age were assumed to likely include all patients in a general hospital setting and were grouped with adult populations. This paper reports the results of all studies investigating CCDS systems studied in adults or populations with an unspecified age.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart demonstrating the study selection process. LILACS: Latin American and Caribbean Health Science Literature; PQDT: ProQuest Dissertations and Theses Global.



Data Abstraction

The data charting form used was iteratively designed based on the study aims. The form was piloted by a single reviewer (KA) and double-checked by a second (JB). Changes to the form were made following discussion between 3 reviewers (KA, JB, and LL). Data charting was performed by 1 reviewer (KA) with ongoing consultation with the review team.

The final data charting form included the components listed in our protocol [26], with minor adjustments as reported in Multimedia Appendix 3 [33-39]. Notably, an additional category *clarity of outcome reporting* was added to the form to account for the variability in outcome reporting transparency. Studies were categorized as having *good* clarity of outcome reporting if they specified the primary outcomes, the outcome analysis method, and the outcome measure definitions and *poor* clarity if the outcomes were not clearly described or there was a substantial reporting discrepancy between the methods and the results. Studies were categorized as having *average* clarity if they fulfilled some criteria of both good and poor.

We accepted any definition of charted data items as specified by the studies. For example, we accepted any definition of systemic inflammatory response syndrome (SIRS), any definition of sepsis, or any cost outcomes specified for the CCDS system. We defined the usability outcome category to follow the ISO definition of usability from ISO 9241-11:2018, section 3.1.1: “extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specific context of use” [33]. We required usability outcomes to be specifically investigated from the perspective of end point users (ie, clinicians). To match this definition change, usability outcomes

were retrospectively categorized into the effectiveness, efficiency, or satisfaction of the CCDS system from the user's perspective.

Analyzing and Reporting the Results

The results were analyzed through both a narrative review and quantitative descriptive analysis. A narrative summary of the data is presented, organized by our 3 aims. The data charted for each aim are summarized into tables using frequency counts and percentages. Graphical figures were also produced, where appropriate.

Owing to the extensive scope of the data charted, several subgroups were collapsed into larger groups to avoid issues of small cell size and to allow for a more meaningful summary. A complete list of the smaller subgroups condensed into the larger groups, organized by table and figure, can be found in Multimedia Appendix 4. Of note, nurses were frequently reported as CCDS system responding personnel and so were grouped separately from other clinicians to better highlight this.

Ethics

Ethical approval or consent to participate was not required for the scoping review. The data were charted from published studies, and no individual information was included.

Results

Study Characteristics

Our initial search identified 22,190 studies, with 12,139 remaining after duplicate removal. Following title, abstract, and full-text screening, 149 studies met our inclusion criteria (Figure 1). Hand searching identified 10 additional studies, resulting in a total of 159 included studies. Of these 159 studies, 124 investigated adult or unspecified populations and were included in this manuscript (Figure 1). A table detailing the main study characteristics for all 124 included studies can be found in Multimedia Appendix 5 [40-163]. In total, 52.4% (65/124) of the studies were categorized as journal articles, 43.5% (54/124) as conference abstracts, and 4% (5/124) as theses (Multimedia Appendices 4 and 5).

Aim 1: Study Context and Design

The context and design characteristics of the studies included in this review are presented in Table 1. Of the 124 included studies, 111 (89.5%) used purely quantitative methods to evaluate CCDS systems (Table 1). Most studies (96/124, 77.4%) used either single cohort (54/124, 43.5%) or before-after (42/124, 33.9%) study designs (Table 1). Very few studies used more robust study designs, such as randomized controlled trials (5/124, 4%), controlled studies (7/124, 5.6%), or interrupted time series (4/124, 3.2%; Table 1). None of the studies reported the use of reporting guidelines. An approximately even distribution of studies was observed across different hospital settings, such as hospital-wide, and specific settings (eg, intensive care unit [ICU], emergency department, and inpatient wards; Table 1).

All studies but 1 (123/124, 99.2%) were published from 2009 onwards, and of the journal articles, 85% (55/65) were published in 2014 or later (Figure 2). Overall, the number of journal articles published steadily increased over time. Of the 65 journal articles, 46 (71%) reported studies conducted in the United States; 2 (3%) each in Germany, Saudi Arabia, and the United Kingdom; 1 (2%) each in Australia, Brazil, Israel, and South Korea; and 9 (14%) did not report which country they were conducted in (Multimedia Appendix 6).

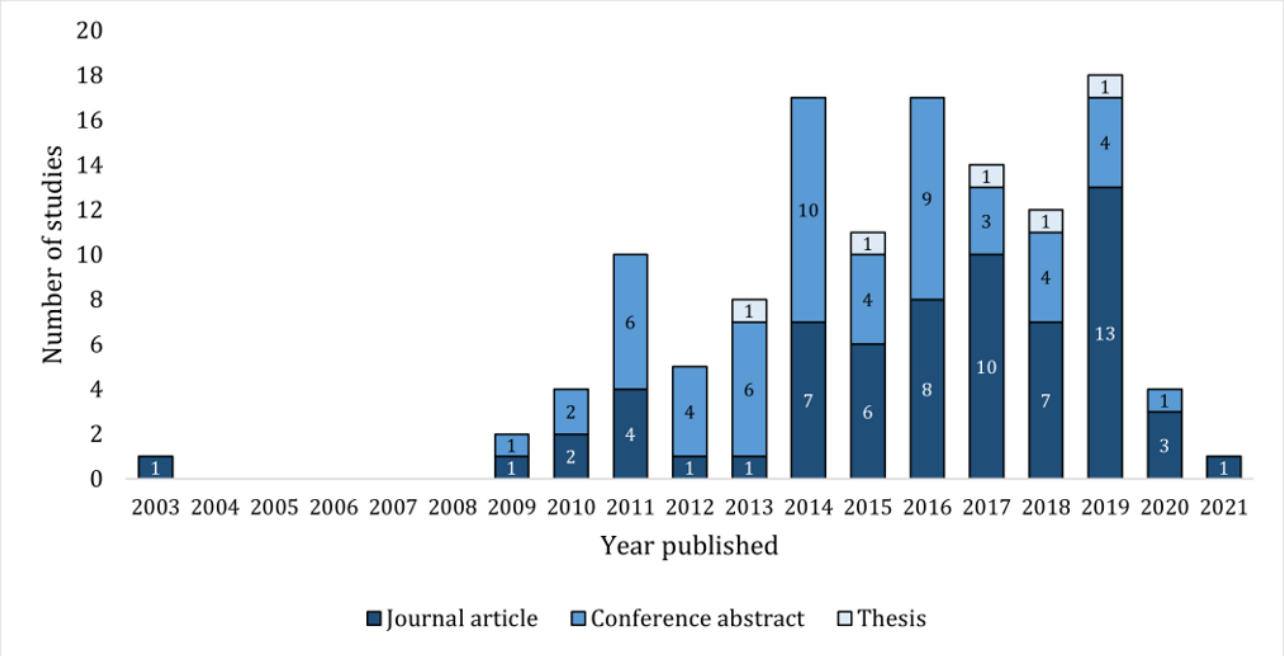
Just over half (66/124, 53.2%) of the studies specified the age of the population as adult. Within these studies, there was a reasonable variation in the actual age range provided. Almost half (29/66, 44%) reported an adult population aged ≥18 years, whereas 30% (20/66) of the studies did not specify an age range further than *adult*. The remaining studies reported populations using thresholds such as aged >14 (1/66, 2%), ≥14 (3/66, 5%), >16 (2/66, 3%), ≥16 (3/66, 5%), ≥19 (6/66, 9%), and ≥70 (1/66, 2%) years, with 2% (1/66) of the studies inconsistently listing multiple thresholds.

Table 1. Study context and design.

Study characteristics	Studies, n (%)			Total (N=124), n (%)
	Conference abstract (n=54)	Journal article (n=65)	Thesis (n=5)	
Method				
Quantitative	50 (92.6)	56 (86.2)	5 (100)	111 (89.5)
Qualitative	1 (1.9)	5 (7.7)	0 (0)	6 (4.8)
Mixed methods	3 (5.6)	4 (6.2)	0 (0)	7 (5.6)
Principal study type				
Surveys or focus groups or heuristics	1 (1.9)	5 (7.7)	0 (0)	6 (4.8)
Case control	1 (1.9)	0 (0)	0 (0)	1 (0.8)
Single cohort	30 (55.6)	22 (33.8)	2 (40)	54 (43.5)
Before and after	13 (24.1)	27 (41.5)	2 (40)	42 (33.9)
Interrupted time series	0 (0)	3 (4.6)	1 (20)	4 (3.2)
Controlled study	3 (5.6)	4 (6.2)	0 (0)	7 (5.6)
Randomized controlled trial	2 (3.7)	3 (4.6)	0 (0)	5 (4)
Insufficient information to determine	4 (7.4)	1 (1.5)	0 (0)	5 (4)
Setting				
Hospital-wide[a]	9 (16.7)	17 (26.2)	1 (20)	27 (21.8)
Intensive care unit	11 (20.4)	14 (21.5)	1 (20)	26 (21)
Emergency department	18 (33.3)	16 (24.6)	2 (40)	36 (29)
Inpatient wards	11 (20.4)	12 (18.5)	1 (20)	24 (19.4)
Specific ward	5 (9.3)	6 (9.2)	0 (0)	11 (8.9)
Number of sites				
1	32 (59.3)	37 (56.9)	2 (40)	71 (57.3)
2-5	6 (11.1)	16 (24.6)	3 (60)	25 (20.2)
>5	0 (0)	6 (9.2)	0 (0)	6 (4.8)
Unspecified	16 (29.6)	6 (9.2)	0 (0)	22 (17.7)
Age group specified?				
Yes	18 (33.3)	44 (67.7)	4 (80)	66 (53.2)
No	36 (66.7)	21 (32.3)	1 (20)	58 (46.8)
Number of participants				
<100	3 (5.6)	8 (12.3)	0 (0)	11 (8.9)
101-500	11 (20.4)	14 (21.5)	1 (20)	26 (21)
501-1000	6 (11.1)	4 (6.2)	0 (0)	10 (8.1)
1001-10,000	9 (16.7)	12 (18.5)	1 (20)	22 (17.7)
>10,001	10 (18.5)	15(23.1)	2 (40)	27 (21.8)
Unspecified	15 (27.8)	12 (18.5)	1 (20)	28 (22.6)
Funding				
Yes	3 (5.6)	26 (40)	1 (20)	30 (24.2)
No	2 (3.7)	13 (20)	0 (0)	15 (12.1)
Unspecified	49 (90.7)	26 (40)	4 (80)	79 (63.7)

[a]If the study setting was not explicitly stated, it was assumed to be hospital-wide.

Figure 2. Number of studies by publication type and year published. Studies published in 2020 include those until September 2020. Studies published in 2021 were found through hand searching.



Aim 2: Study Outcomes

The outcomes investigated by the included journal articles and the conference abstracts and theses are presented in Table 2 and Multimedia Appendix 7, respectively. Of the 4 predefined outcome categories, patient outcomes were reported in the highest number of studies (107/124, 86.3%; Figure 3). Sepsis treatment and management outcomes were reported in 60.5% (75/124) of the studies, CCDS system usability outcomes in 11.3% (14/124), and cost outcomes in 7.3% (9/124; Figure 3).

Overall, only 31.5% (39/124) of the studies had good clarity in outcome reporting (Figure 4). Generally, studies had average (62/124, 50%) or poor clarity (23/124, 18.5%). Unsurprisingly, journal articles had better clarity of outcome reporting, with 40% (26/65) of the articles having good clarity, compared with 22% (13/59) of the conference abstracts or theses (Figure 4).

In the 65 journal articles, mortality was the most frequently reported patient outcome (39/65, 60%). Overall, 35 different types of mortality measures were reported 55 times across 39 studies (Multimedia Appendix 8). Of these, in-hospital mortality was the most frequently reported (13/55, 24%; Multimedia Appendix 8). Sepsis identification, length of stay, and *other* patient outcomes were also frequently reported, appearing in 38% (25/65), 34% (22/65), and 35% (23/65) of the articles, respectively (Table 2; see Multimedia Appendix 4 for the expanded list of included outcomes). ICU admission was the least reported patient outcome (12/65, 18%). In the sepsis treatment and management outcome category, antibiotic-related and *other* were the most frequently reported outcomes in journal articles (27/65, 42% and 31/65, 48%, respectively), followed by lactate-, fluids-, and blood culture–related outcomes (17/65, 26%; 14/65, 22%; and 14/65, 22%, respectively; Table 2; see Multimedia Appendix 4 for expanded list of included outcomes). Overall, sepsis bundle or protocol compliance was the least reported outcome in journal articles (12/65, 18%).

CCDS system usability outcomes were reported in similar numbers of journal articles, with 12% (8/65) of the journal articles reporting on the efficiency of the system, 11% (7/65) on system effectiveness, and 11% (7/65) reporting on users' satisfaction with the system (Table 2). Among the CCDS system-related cost outcomes, cost was reported in the greatest number of journal articles (5/65, 8%), whereas cost-effectiveness or savings were reported in only 5% (3/65) of the articles (Table 2).

Table 2. Main outcomes and outcome categories in journal articles.

Outcome categories	Outcome classification[a], n (%[b])			Total, (n=65), n (%)[c]
	Primary	Secondary	Not specified[d]	
Patient outcomes				
Mortality	11 (28)	9 (23)	19 (49)	39 (60)
Sepsis identification	10 (40)	3 (12)	12 (48)	25 (38)
Length of stay	3 (14)	10 (45)	9 (41)	22 (34)
Intensive care unit admission	0 (0)	5 (42)	7 (58)	12 (18)
Other	4 (17)	8 (35)	11 (48)	23 (35)
Sepsis treatment and management				
Antibiotics	6 (22)	9 (33)	12 (44)	27 (42)
Lactate	2 (12)	6 (35)	9 (53)	17 (26)
Fluids	2 (14)	4 (29)	8 (57)	14 (22)
Blood culture	2 (14)	4 (29)	8 (57)	14 (22)
Sepsis bundle or protocol compliance	2 (17)	4 (33)	6 (50)	12 (18)
Other	5 (16)	7 (23)	19 (61)	31 (48)
Usability				
Efficiency	0 (0)	2 (25)	6 (75)	8 (12)
Effectiveness	0 (0)	1 (14)	6 (86)	7 (11)
Satisfaction	0 (0)	3 (43)	4 (57)	7 (11)
Cost				
Cost	1 (20)	2 (40)	2 (40)	5 (8)
Cost-effectiveness or savings	0 (0)	1 (33)	2 (67)	3 (5)

[a]Some studies reported both primary, secondary, or nonspecified outcomes within the same outcome group. To avoid double-counting these studies, secondary outcomes were not counted in favor of counting primary outcomes. Similarly, nonspecified outcomes were not counted in favor of primary or secondary outcomes. For example, a study may have the primary outcome mortality (30-day) and the secondary outcome mortality (7-day), which would both fall into the mortality outcome group. In this example the study would be counted as having mortality as the primary outcome.

[b]These percentages were calculated as row percentages, that is, using the number in the "Total" column in each row as the denominator.

[c]The percentages were calculated from the number of journal articles (n=65), not the number of total outcomes. As many journal articles reported multiple outcomes, there were more than 65 outcomes in each category, and therefore, the percentages will add up to more than 100%.

[d]The study did not specify whether the outcome was primary or secondary.

Figure 3. Proportion of studies reporting each outcome category.

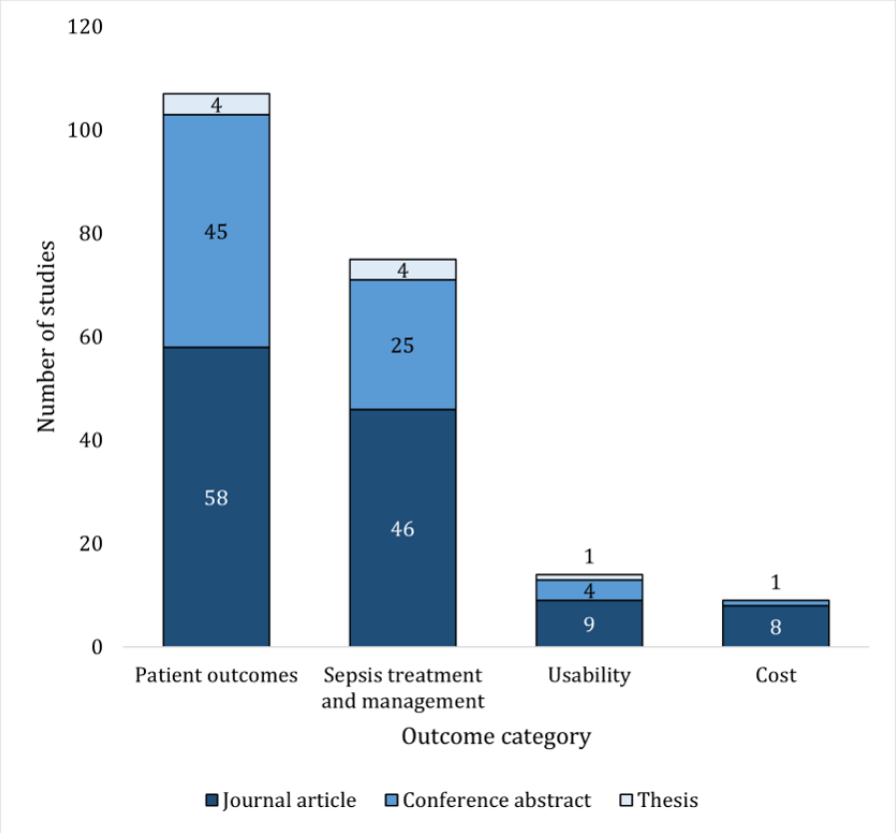
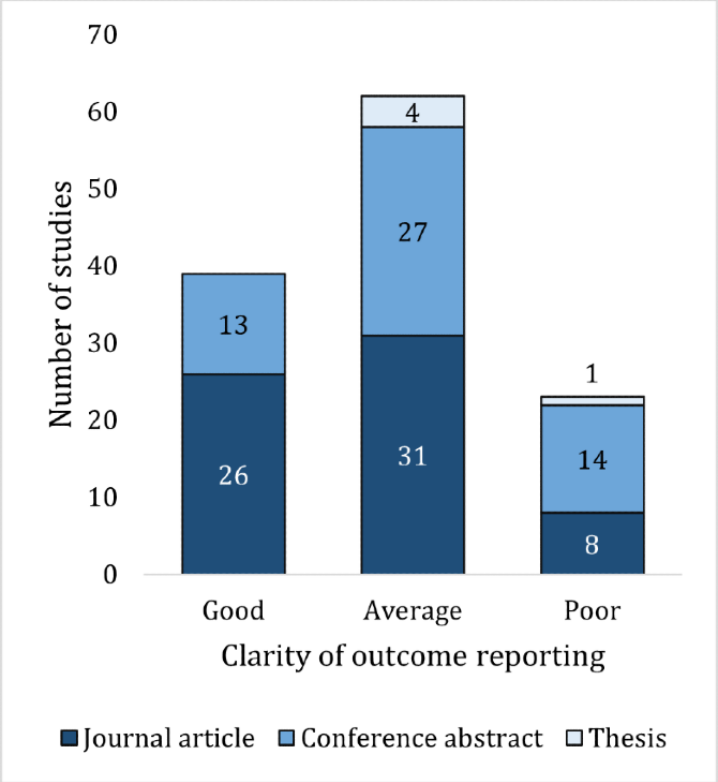


Figure 4. Clarity of outcome reporting in the studies.



Aim 3: CCDS Characteristics

The characteristics of the CCDS systems reported in the included studies are presented in Table 3. Half (64/124, 51.6%) of the studies, most of which were journal articles (44/64, 69%), implemented homegrown CCDS systems. Of the 124 studies, only 13 (10.5%), including 10 (77%) journal articles, implemented commercial CCDS systems, of which 69% (9/13) were the St John's Sepsis Surveillance Agent (Cerner Corporation, Kansas City, Missouri, United States; Table 3).

Most included studies (95/124, 76.6%) evaluated *live* CCDS systems only, where the CCDS was implemented and actively sending alerts. Silent CCDS, where the system would run in real time but not send clinical alerts, were implemented by 7.3% (9/124) of studies, and 11.3% (14/124) of studies implemented both silent and live CCDS, either sequentially or concurrently.

SIRS alone was the most frequently used CCDS clinical criteria for sepsis identification (50/124, 40.3%), followed by SIRS combined with organ dysfunction (28/124, 22.6%), and SIRS combined with other criteria (23/124, 18.5%; Table 3). In addition, a diverse range of other criteria were used by 32.3% (40/124) of studies (Multimedia Appendix 4), while 7.3% (9/124) did not specify the clinical criteria used (Table 3)

Over half of the studies reported the implementation of CCDS systems alongside numerous other related interventions (68/124, 54.8%), such as staff education programs and antibiotic order sets (Table 3). The most common type of concurrent intervention used was clinical protocols in 41.9% (52/124) of the studies (Table 3).

Most commonly, studies reported nurses (51/124, 41.1%) or other clinicians (37/124, 29.8%) as the main CCDS alert responding personnel (Table 3). Some studies reported on CCDS with response teams (12/124, 9.7%), study coordinators (8/124, 6.5%), or other personnel (11/124, 8.9%) responding to the alerts. Of the 124 studies, 33 (26.6%) reported the use of the electronic health record to distribute CCDS alerts, 26 (21%) the use of pagers, 15 (12.1%) the use of a patient dashboard or work list, and 8 (6.5%) the use of another form of alert delivery (Table 3).

Table 3. Computerized clinical decision support (CCDS) characteristics.

CCDS characteristic	Studies, n (%[a])			Total (N=124), n (%[a])
	Conference abstract (n=54)	Journal article (n=65)	Thesis (n=5)	
CCDS type				
Homegrown	18 (33.3)	44 (67.7)	2 (40)	64 (51.6)
Commercial	3 (5.6)	10 (15.4)	0 (0)	13 (10.5)
St John sepsis (Cerner Corporation, Kansas City, Missouri, United States)	2 (3.7)	7 (10.7)	0 (0)	9 (7.3)
PREDEC ALARM (Löser Medizintechnik GmbH, Leipzig, Germany)	0 (0)	1 (1.5)	0 (0)	1 (0.8)
Unspecified	1 (1.9)	2 (3.1)	0 (0)	3 (2.4)
Unspecified	33 (61.1)	11 (16.9)	3 (60)	47 (37.9)
Silent or live?				
Live	41 (75.8)	49 (75.4)	5 (100)	95 (76.6)
Silent	5 (9.3)	4 (6.2)	0 (0)	9 (7.3)
Both	4 (7.4)	10 (15.4)	0 (0)	14 (11.3)
Unspecified	4 (7.4)	2 (3.1)	0 (0)	6 (4.8)
CCDS criteria				
SIRS[b]	24 (44.4)	24 (36.9)	2 (40)	50 (40.3)
SIRS + organ dysfunction	11 (20.4)	17 (26.2)	0 (0)	28 (22.6)
SIRS + other	9 (16.7)	12 (18.5)	2 (40)	23 (18.5)
Other	17 (31.5)	23 (35.4)	0 (0)	40 (32.3)
Unspecified	3 (5.6)	5 (7.7)	1 (20)	9 (7.3)
Related interventions				
Clinical protocol	16 (29.6)	34 (52.3)	2 (40)	52 (41.9)
Education and staff resources	8 (14.8)	25 (38.5)	1 (20)	34 (27.4)
Electronic or infrastructure changes	2 (3.7)	6 (9.2)	0 (0)	8 (6.5)
Response or leadership team	4 (7.4)	17 (26.2)	1 (20)	22 (17.7)
Order sets	10 (18.5)	17 (26.2)	0 (0)	27 (21.8)
Feedback	0 (0)	10 (15.4)	0 (0)	10 (8.1)
None	32 (59.3)	22 (33.8)	2 (40)	56 (45.2)
Responding personnel				
Nurses[c]	12 (22.2)	38 (58.5)	1 (20)	51 (41.1)
Other clinicians	11 (20.4)	26 (40)	0 (0)	37 (29.8)
Response team	4 (7.4)	8 (12.3)	0 (0)	12 (9.7)
Study coordinator	0 (0)	8 (12.3)	0 (0)	8 (6.5)
Other	4 (7.4)	6 (9.2)	1 (20)	11 (8.9)
Unspecified	27 (50)	7 (10.8)	3 (60)	37 (29.8)
Alert delivery				
Electronic patient record	6 (11.1)	27 (41.5)	0 (0)	33 (26.6)
Pager	4 (7.4)	20 (30.8)	2 (40)	26 (21.0)
Patient dashboard or working list	4 (7.4)	10 (15.4)	1 (20)	15 (12.1)
Other	1 (1.9)	6 (9.2)	1 (20)	8 (6.5)

CCDS characteristic	Studies, n (%[a])			Total (N=124), n (%[a])
	Conference abstract (n=54)	Journal article (n=65)	Thesis (n=5)	
Unspecified	42 (77.8)	19 (29.2)	2 (40)	63 (50.8)

[a]As some studies reported multiple characteristics within each category, there were more than the total number of studies, and therefore, the percentages may add up to more than 100%.

[b]SIRS: systemic inflammatory response syndrome.

[c]As nurses are frequently reported as CCDS system responding personnel, they were grouped separately from other clinicians.

Discussion

Principal Findings

This review canvassed 124 studies in total, representing a comprehensive overview of current research, including an extensive body of gray literature. Over half of the included studies were journal articles (65/124, 52.4%), and nearly all studies were published in the last decade, indicating the considerable volume of recent research investigating the use of CCDS systems for early detection of adult inpatients with sepsis. Our findings demonstrate the substantial diversity of studies across all three aims: (1) the context and design of the study, (2) the type and measurement of outcomes investigated, and (3) the design and implementation of the CCDS system evaluated. We identified little research into the effects of CCDS on patient morbidity or CCDS usability and cost outcomes, highlighting key knowledge gaps in the literature. Our review also underlines the need for robust study designs, as well as improved generalizability and reporting in future studies.

Variability Across Studies

There is extensive heterogeneity in the current literature investigating the implementation and evaluation of CCDS systems for early sepsis detection in adult hospital patients. In particular, there was considerable diversity displayed in the chosen clinical criteria for sepsis identification across the studies included in our review. Although many studies used the SIRS criteria, alone or with adjuncts, there was a substantial range of other criteria used (Multimedia Appendix 4). This can be attributed to the extremely diverse presentations of patients with sepsis, which has led to the development of numerous different clinical scores for sepsis detection [15-18,164]. In addition, our findings demonstrated variability in the method of alert delivery, personnel who respond to alerts, and concurrent implementation of related interventions. Studies were conducted across a range of different hospital settings, including hospital-wide or specific sites, such as the emergency department or ICU (Table 1). The chosen threshold for what age participants were included in the study was also quite variable, with studies defining their *adult* population using cutoff points ranging from 14 to 19 years and older. Finally, our review illuminated the expansive number of outcomes used to evaluate and investigate sepsis CCDS systems. Previous systematic reviews have similarly highlighted this diversity [13,21,22,165,166]. This heterogeneity across settings, participants, CCDS system characteristics, and outcomes makes it difficult to compare studies and to make general statements regarding sepsis CCDS systems.

This diversity can be partially attributed to the novel nature of sepsis CCDS systems and the recent emergence of the field. Our findings show a vast expansion of the literature, with three-quarters of studies published since 2014 (Figure 2). Owing to this recent rapid development of the field and the simultaneous evolution of information and communication technology in health care [21,167], there is no well-established research strategy or dogma for this specific area. Consequently, different authors have designed and executed their studies using a diverse range of variables and study design methodologies.

This variability can also be attributed to the complexity involved in the implementation of health care interventions [168,169]. To characterize this complexity, Greenhalgh et al [170] have designed the NASSS (nonadoption, abandonment, scale-up, spread, and sustainability) framework. In the case of CCDS systems for early sepsis detection in hospitals, the 7 NASSS framework domains can be identified as follows: sepsis (the condition); the CCDS system (the technology); the commercial and health-associated value of CCDS systems (value proposition and value chain); the responding personnel (the adopters); the hospital setting (the organization); the local, state, or national health system (the wider system); and software plasticity (embedding and adaptation over time). Our findings demonstrate that these domains are extremely diverse across the included studies, presenting many variables and variable combinations, consequentially expanding the complexity involved in sepsis CCDS system implementation. As the complexity of a system has been associated with its capacity for successful and sustainable implementation [25], this heterogeneity could detrimentally impact the performance of sepsis CCDS systems. To counter this issue, Greenhalgh et al [25] highlights the importance of system usability and adaptability, suggesting that a user-centered and iterative approach is needed, centralizing the involvement of relevant providers in the implementation plan. Unfortunately, our findings indicate that few of the included studies investigated the usability of sepsis CCDS systems.

Knowledge Gaps for Future Research

Patient Outcomes

Although patient outcomes were the most commonly reported outcome (Figure 3), none of the included studies directly measured the effect of CCDS systems on sepsis morbidity. Surviving sepsis is associated with cognitive impairment, higher mortality rates across the life span, physical disability, and mental health issues [3,4,6,7,9]. This not only substantially reduces the quality of life of survivors of sepsis but also presents an enormous financial burden on both patients and health care

systems [5,8,171,172]. Reducing sepsis morbidity rates through CCDS use would be extremely valuable for patient health and quality of life and in mitigating personal and health care–related costs. Consequently, it is highlighted as a clear gap in the evidence base.

Usability and Cost Outcomes

We identified inadequate investigation of CCDS-related usability and cost outcomes, with most included studies focusing on clinical outcomes. The ability of a user to successfully operate a clinical information system is critical to the success of a system [25,173-175]. This is accentuated in the busy hospital environment, where medical providers are often time poor and carry enormous mental burdens [21,42]. Of particular concern in sepsis CCDS systems is the occurrence of alert fatigue [176,177]. Alert fatigue refers to when clinicians become desensitized to clinical alerts and consequently ignore or turn off alarm systems, potentially missing real sepsis cases [176,177]. This can have serious implications for patient outcomes. Strategies to ensure good CCDS system usability include incorporating human factor design elements, integrating CCDS system sepsis workflows into current medical emergency clinical pathways, and linking CCDS systems with existing clinical deterioration policies [42,178-180]. Only 11.3% (14/124) of the studies we investigated included usability outcomes, with only 14% (2/14) of these studies [70,93] evaluating alert fatigue. This represents a clear gap in the current literature for further research to support the successful implementation of appropriate, usable, and effective CCDS systems for early sepsis detection in hospitals.

In addition, very few studies investigated cost outcomes of CCDS system implementation. Sepsis is an extremely expensive condition to treat [171]. It has been reported to cost more than US $20 billion annually and is listed as the most financially costly condition in US hospitals [172]. Sepsis-related costs can range from extensive hospital costs during acute treatment to high long-term treatment and rehabilitation costs in survivors of sepsis [3,171,181]. Determining the cost-effectiveness of sepsis CCDS systems would assist in establishing the financial feasibility of implementation in hospitals and support widespread implementation.

Study Design and Generalizability

Few studies applied robust study designs such as randomized controlled trials, interrupted time series, stepped wedge clusters, and controlled trials. Future research in this area should attempt to use more rigorous methodology to present stronger evidence.

Approximately three-quarters of the included journal articles were conducted in the United States (Multimedia Appendix 6), limiting generalizability to other settings. A recent study demonstrated that the bulk of the sepsis burden is in countries with a low, low-middle, or middle sociodemographic index [2]. Future studies investigating the use of CCDS systems for adult sepsis inpatient identification should be encouraged to examine trends in countries outside the United States. In particular, CCDS systems should be evaluated in low- to middle-income countries when possible, given the limited availability of electronic health care technology in such regions.

Reporting and Transparency

A large proportion of studies did not specify important study design, CCDS system, and main outcome details (Tables 1-3). An unexpectedly high number of journal articles did not report these details nor did most conference abstracts; however, this is more understandable given word limit constraints. Of particular concern is that almost two-thirds of the included journal articles were found to have an average or poor clarity of outcome reporting (Figure 4). None of the studies included in this review published the use of reporting guidelines, likely because of many journals not specifically requiring it. Overall, we found that the quality of reporting is low and identified a need for improved reporting and transparency throughout the literature.

Strengths and Limitations

This scoping review comprehensively canvassed the literature investigating knowledge-based implemented CCDS systems for early sepsis detection in adult hospital patients. Its strength lies in this breadth of coverage and the wide range of study elements examined. The review followed the PRISMA-ScR expansion [28] guidelines, the Joanna Briggs Institute Reviewer's Manual [27], and the framework presented by Arksey and O'Malley [29].

A limitation of this scoping review is that it only included studies written in English or had English translations readily available. Furthermore, only a sample of the charted data was double-checked by a second reviewer, potentially resulting in a higher error margin. However, the data charting forms were well structured, and any issues occurring during charting were fully discussed among the research team to reach consensus.

Conclusions

This review highlights the extensive variability in the design, outcomes, and system characteristics in studies investigating the use of CCDS for the early detection of sepsis in adult inpatients. This heterogeneity can be largely attributed to the considerable complexity of sepsis, CCDS software, and the hospital environment. Our findings have identified clear gaps in the current literature, with few studies investigating CCDS system usability, cost, or the effects on patient morbidity. There are limited studies conducted outside the United States or with robust study designs. Our findings have illustrated frequent poor reporting of CCDS system information and study outcomes. It is critically important for future research to close these knowledge gaps, ensuring comprehensive evaluation of these rapidly emerging sepsis CCDS systems.

Acknowledgments

The authors would like to thank Mr Jeremy Cullis, an experienced clinical librarian, for his advice and expertise in designing the final search strategy and translating it to other databases.

The authors received no financial support or funding for this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist.
[PDF File (Adobe PDF File), 507 KB - jmir_v24i2e31083_app1.pdf]

Multimedia Appendix 2
Final MEDLINE search strategy.
[PDF File (Adobe PDF File), 75 KB - jmir_v24i2e31083_app2.pdf]

Multimedia Appendix 3
Adjustments made to data charting form.
[PDF File (Adobe PDF File), 374 KB - jmir_v24i2e31083_app3.pdf]

Multimedia Appendix 4
Definitions of groups combining multiple subgroups.
[PDF File (Adobe PDF File), 196 KB - jmir_v24i2e31083_app4.pdf]

Multimedia Appendix 5
Main study characteristics.
[PDF File (Adobe PDF File), 311 KB - jmir_v24i2e31083_app5.pdf]

Multimedia Appendix 6
Number of journal articles by country.
[PDF File (Adobe PDF File), 71 KB - jmir_v24i2e31083_app6.pdf]

Multimedia Appendix 7
Main outcomes and outcome categories in gray literature.
[PDF File (Adobe PDF File), 209 KB - jmir_v24i2e31083_app7.pdf]

Multimedia Appendix 8
Types of mortality reported in journal articles.
[PDF File (Adobe PDF File), 101 KB - jmir_v24i2e31083_app8.pdf]

References

1. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, et al. The third international consensus definitions for sepsis and septic shock (Sepsis-3). J Am Med Assoc 2016 Feb 23;315(8):801-810. [doi: 10.1001/jama.2016.0287] [Medline: 26903338]
2. Rudd KE, Johnson SC, Agesa KM, Shackelford KA, Tsoi D, Kievlan DR, et al. Global, regional, and national sepsis incidence and mortality, 1990-2017: analysis for the Global Burden of Disease Study. Lancet 2020 Jan 18;395(10219):200-211 [FREE Full text] [doi: 10.1016/S0140-6736(19)32989-7] [Medline: 31954465]
3. Iwashyna TJ, Ely EW, Smith DM, Langa KM. Long-term cognitive impairment and functional disability among survivors of severe sepsis. J Am Med Assoc 2010 Oct 27;304(16):1787-1794 [FREE Full text] [doi: 10.1001/jama.2010.1553] [Medline: 20978258]
4. Winters BD, Eberlein M, Leung J, Needham DM, Pronovost PJ, Sevransky JE. Long-term mortality and quality of life in sepsis: a systematic review. Crit Care Med 2010 May;38(5):1276-1283. [doi: 10.1097/CCM.0b013e3181d8cc1d] [Medline: 20308885]

5. Yende S, Austin S, Rhodes A, Finfer S, Opal S, Thompson T, et al. Long-term quality of life among survivors of severe sepsis: analyses of two international trials. Crit Care Med 2016 Aug;44(8):1461-1467 [FREE Full text] [doi: 10.1097/CCM.0000000000001658] [Medline: 26992066]
6. Davydow DS, Hough CL, Langa KM, Iwashyna TJ. Symptoms of depression in survivors of severe sepsis: a prospective cohort study of older Americans. Am J Geriatr Psychiatry 2013 Sep;21(9):887-897 [FREE Full text] [doi: 10.1016/j.jagp.2013.01.017] [Medline: 23567391]
7. Barichello T, Sayana P, Giridharan VV, Arumanayagam AS, Narendran B, Giustina AD, et al. Long-term cognitive outcomes after sepsis: a translational systematic review. Mol Neurobiol 2019 Jan;56(1):186-251. [doi: 10.1007/s12035-018-1048-2] [Medline: 29687346]
8. Alam N, Panday RS, Heijnen JR, van Galen LS, Kramer MH, Nanayakkara PW. Long-term health related quality of life in patients with sepsis after intensive care stay: a systematic review. Acute Med 2017;16(4):164-169. [Medline: 29300794]
9. Angus DC, van der Poll T. Severe sepsis and septic shock. N Engl J Med 2013 Aug 29;369(9):840-851. [doi: 10.1056/NEJMra1208623] [Medline: 23984731]
10. Seymour CW, Gesten F, Prescott HC, Friedrich ME, Iwashyna TJ, Phillips GS, et al. Time to treatment and mortality during mandated emergency care for sepsis. N Engl J Med 2017 Jun 08;376(23):2235-2244 [FREE Full text] [doi: 10.1056/NEJMoa1703058] [Medline: 28528569]
11. Paoli CJ, Reynolds MA, Sinha M, Gitlin M, Crouser E. Epidemiology and costs of sepsis in the United States-an analysis based on timing of diagnosis and severity level. Crit Care Med 2018 Dec;46(12):1889-1897 [FREE Full text] [doi: 10.1097/CCM.0000000000003342] [Medline: 30048332]
12. Rhodes A, Evans LE, Alhazzani W, Levy MM, Antonelli M, Ferrer R, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock: 2016. Intensive Care Med 2017 Mar;43(3):304-377. [doi: 10.1007/s00134-017-4683-6] [Medline: 28101605]
13. Joshi M, Ashrafian H, Arora S, Khan S, Cooke G, Darzi A. Digital alerting and outcomes in patients with sepsis: systematic review and meta-analysis. J Med Internet Res 2019 Dec 20;21(12):e15166 [FREE Full text] [doi: 10.2196/15166] [Medline: 31859672]
14. Cecconi M, Evans L, Levy M, Rhodes A. Sepsis and septic shock. Lancet 2018 Jul 07;392(10141):75-87. [doi: 10.1016/S0140-6736(18)30696-2] [Medline: 29937192]
15. McLymont N, Glover GW. Scoring systems for the characterization of sepsis and associated outcomes. Ann Transl Med 2016 Dec;4(24):527 [FREE Full text] [doi: 10.21037/atm.2016.12.53] [Medline: 28149888]
16. Churpek MM, Snyder A, Han X, Sokol S, Pettit N, Howell MD, et al. Quick sepsis-related organ failure assessment, systemic inflammatory response syndrome, and early warning scores for detecting clinical deterioration in infected patients outside the intensive care unit. Am J Respir Crit Care Med 2017 Apr 01;195(7):906-911 [FREE Full text] [doi: 10.1164/rccm.201604-0854OC] [Medline: 27649072]
17. Seymour CW, Liu VX, Iwashyna TJ, Brunkhorst FM, Rea TD, Scherag A, et al. Assessment of clinical criteria for sepsis: for the third international consensus definitions for sepsis and septic shock (Sepsis-3). J Am Med Assoc 2016 Feb 23;315(8):762-774. [doi: 10.1001/jama.2016.0288] [Medline: 26903335]
18. Li L, Rathnayake K, Green M, Shetty A, Fullick M, Walter S, et al. Comparison of the quick sepsis-related organ failure assessment (qSOFA) and adult sepsis pathway in predicting adverse outcomes among adult patients on general wards: a retrospective observational cohort study. Intern Med J 2020 Jan 06. [doi: 10.1111/imj.14746] [Medline: 31908090]
19. Bhattacharjee P, Edelson DP, Churpek MM. Identifying patients with sepsis on the hospital wards. Chest 2017 Apr;151(4):898-907 [FREE Full text] [doi: 10.1016/j.chest.2016.06.020] [Medline: 27374948]
20. Marik PE. Don't miss the diagnosis of sepsis!. Crit Care 2014 Sep 27;18(5):529 [FREE Full text] [doi: 10.1186/s13054-014-0529-6] [Medline: 25675360]
21. Wulff A, Montag S, Marschollek M, Jack T. Clinical decision-support systems for detection of systemic inflammatory response syndrome, sepsis, and septic shock in critically ill patients: a systematic review. Methods Inf Med 2019 Dec;58(S 02):43-57 [FREE Full text] [doi: 10.1055/s-0039-1695717] [Medline: 31499571]
22. Makam AN, Nguyen OK, Auerbach AD. Diagnostic accuracy and effectiveness of automated electronic sepsis alert systems: a systematic review. J Hosp Med 2015 Jun;10(6):396-402 [FREE Full text] [Medline: 25758641]
23. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digit Med 2020;3:17 [FREE Full text] [doi: 10.1038/s41746-020-0221-y] [Medline: 32047862]
24. Petersen C, Smith J, Freimuth RR, Goodman KW, Jackson GP, Kannry J, et al. Recommendations for the safe, effective use of adaptive CDS in the US healthcare system: an AMIA position paper. J Am Med Inform Assoc 2021 Mar 18;28(4):677-684. [doi: 10.1093/jamia/ocaa319] [Medline: 33447854]
25. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Analysing the role of complexity in explaining the fortunes of technology programmes: empirical application of the NASSS framework. BMC Med 2018 May 14;16(1):66 [FREE Full text] [doi: 10.1186/s12916-018-1050-6] [Medline: 29754584]

26. Li L, Ackermann K, Baker J, Westbrook J. Use and evaluation of computerized clinical decision support systems for early detection of sepsis in hospitals: protocol for a scoping review. JMIR Res Protoc 2020 Nov 20;9(11):e24899 [FREE Full text] [doi: 10.2196/24899] [Medline: 33215998]
27. Peters M, Godfrey C, McInerney P, Munn Z, Tricco A, Khalil H. Chapter 11: scoping reviews. In: Aromataris E, Munn Z, editors. JBI Manual for Evidence Synthesis. Australia: JBI - The University of Adelaide; 2020.
28. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
29. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
30. Bonet M, Pileggi VN, Rijken MJ, Coomarasamy A, Lissauer D, Souza JP, et al. Towards a consensus definition of maternal sepsis: results of a systematic review and expert consultation. Reprod Health 2017 May 30;14(1):67 [FREE Full text] [doi: 10.1186/s12978-017-0321-6] [Medline: 28558733]
31. Schlapbach LJ. Paediatric sepsis. Curr Opin Infect Dis 2019 Oct;32(5):497-504. [doi: 10.1097/QCO.0000000000000583] [Medline: 31335441]
32. Shane AL, Sánchez PJ, Stoll BJ. Neonatal sepsis. Lancet 2017 Oct 14;390(10104):1770-1780. [doi: 10.1016/S0140-6736(17)31002-4] [Medline: 28434651]
33. Ergonomics of human-system interaction - Part 11: usability: definitions and concepts. ISO standard no. 9241-11:2018(EN). The International Organization for Standardization. 2018. URL: https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en [accessed 2021-02-26]
34. Viswanathan M, Berkman N, Dryden D, Hartling L. Assessing risk of bias and confounding in observational studies of interventions or exposures : further development of the RTI item bank. In: AHRQ Methods for Effective Health Care. Rockville (MD): Agency for Healthcare Research and Quality (US); 2013.
35. Ranganathan P, Aggarwal R. Study designs: Part 1 - An overview and classification. Perspect Clin Res 2018;9(4):184-186 [FREE Full text] [Medline: 30319950]
36. Aggarwal R, Ranganathan P. Study designs: Part 2 - Descriptive studies. Perspect Clin Res 2019;10(1):34-36 [FREE Full text] [doi: 10.4103/picr.PICR_154_18] [Medline: 30834206]
37. Ranganathan P, Aggarwal R. Study designs: Part 3 - Analytical observational studies. Perspect Clin Res 2019;10(2):91-94 [FREE Full text] [doi: 10.4103/picr.PICR_35_19] [Medline: 31008076]
38. Aggarwal R, Ranganathan P. Study designs: Part 4 - Interventional studies. Perspect Clin Res 2019;10(3):137-139 [FREE Full text] [doi: 10.4103/picr.PICR_91_19] [Medline: 31404185]
39. Aggarwal R, Ranganathan P. Study designs: Part 5 - Interventional studies (II). Perspect Clin Res 2019;10(4):183-186 [FREE Full text] [Medline: 31649869]
40. Clinical score usefulness and EHR integration (ScoreInt). ClinicalTrials.gov identifier: NCT02755025. 2017 Mar 29. URL: https://clinicaltrials.gov/ct2/show/NCT02755025 [accessed 2020-08-17]
41. Aakre C, Franco PM, Ferreyra M, Kitson J, Li M, Herasevich V. Prospective validation of a near real-time EHR-integrated automated SOFA score calculator. Int J Med Inform 2017 Jul;103:1-6. [doi: 10.1016/j.ijmedinf.2017.04.001] [Medline: 28550994]
42. Aakre CA, Kitson JE, Li M, Herasevich V. Iterative user interface design for automated sequential organ failure assessment score calculator in sepsis detection. JMIR Hum Factors 2017 May 18;4(2):e14 [FREE Full text] [doi: 10.2196/humanfactors.7567] [Medline: 28526675]
43. Acquah S, Stoller M, Heller M, Krupka M, Wang C, Smith M, et al. Utilizing electronic alerts and IVC ultrasound to improve outcomes for sepsis care in an urban ED. In: Proceedings of the American Thoracic Society International Conference, ATS 2014. 2014 Presented at: American Thoracic Society International Conference, ATS 2014; May 16-21, 2014; San Diego, CA United States URL: https://www.atsjournals.org/doi/pdf/10.1164/ajrccm-conference.2014.189.1_MeetingAbstracts.A2817
44. Afshar M, Arain E, Ye C, Gilbert E, Xie M, Lee J, et al. Patient outcomes and cost-effectiveness of a sepsis care quality improvement program in a health system. Crit Care Med 2019 Oct;47(10):1371-1379 [FREE Full text] [doi: 10.1097/CCM.0000000000003919] [Medline: 31306176]
45. Alsolamy S, Al Salamah M, Al Thagafi M, Al-Dorzi HM, Marini AM, Aljerian N, et al. Diagnostic accuracy of a screening electronic alert tool for severe sepsis and septic shock in the emergency department. BMC Med Inform Decis Mak 2014 Dec 05;14:105 [FREE Full text] [doi: 10.1186/s12911-014-0105-7] [Medline: 25476738]
46. Amland RC, Hahn-Cover KE. Clinical decision support for early recognition of sepsis. Am J Med Qual 2016;31(2):103-110 [FREE Full text] [doi: 10.1177/1062860614557636] [Medline: 25385815]
47. Amland RC, Haley JM, Lyons JJ. A multidisciplinary sepsis program enabled by a two-stage clinical decision support system: factors that influence patient outcomes. Am J Med Qual 2016 Nov;31(6):501-508 [FREE Full text] [doi: 10.1177/1062860615606801] [Medline: 26491116]

48. Amland RC, Sutariya BB. Quick sequential [Sepsis-related] organ failure assessment (qSOFA) and St. John Sepsis Surveillance Agent to detect patients at risk of sepsis: an observational cohort study. Am J Med Qual 2018;33(1):50-57 [FREE Full text] [doi: 10.1177/1062860617692034] [Medline: 28693336]
49. Amland RC, Sutariya BB. An investigation of sepsis surveillance and emergency treatment on patient mortality outcomes: an observational cohort study. JAMIA Open 2018 Jul;1(1):107-114 [FREE Full text] [doi: 10.1093/jamiaopen/ooy013] [Medline: 31984322]
50. Arabi YM, Al-Dorzi HM, Alamry A, Hijazi R, Alsolamy S, Al Salamah M, et al. The impact of a multifaceted intervention including sepsis electronic alert system and sepsis response team on the outcomes of patients with sepsis and septic shock. Ann Intensive Care 2017 Dec;7(1):57 [FREE Full text] [doi: 10.1186/s13613-017-0280-7] [Medline: 28560683]
51. Austrian JS, Jamin CT, Doty GR, Blecker S. Impact of an emergency department electronic sepsis surveillance system on patient mortality and length of stay. J Am Med Inform Assoc 2018 May 01;25(5):523-529 [FREE Full text] [doi: 10.1093/jamia/ocx072] [Medline: 29025165]
52. Bansal V, Festić E, Mangi MA, Decicco NA, Reid AN, Gatch EL, et al. Early machine-human interface around sepsis severity identification: from diagnosis to improved management? Acta Med Acad 2018 May;47(1):27-38 [FREE Full text] [doi: 10.5644/ama2006-124.212] [Medline: 29957969]
53. Becker T, Herres J, Kelly J. Improved identification and management of patients with sepsis, using an electronic screening tool. In: Proceedings of the Annual Meeting of the Society for Academic Emergency Medicine. 2018 Presented at: 2018 Annual Meeting of the Society for Academic Emergency Medicine; May 15-18, 2018; Indianapolis, IN, United States p. S81. [doi: 10.1111/acem.13424]
54. Benson L, Hasenau S, O'Connor N, Burgermeister D. The impact of a nurse practitioner rapid response team on systemic inflammatory response syndrome outcomes. Dimens Crit Care Nurs 2014;33(3):108-115. [doi: 10.1097/DCC.0000000000000046] [Medline: 24704733]
55. Berger T, Birnbaum A, Bijur P, Kuperman G, Gennis P. A computerized alert screening for severe sepsis in emergency department patients increases lactate testing but does not improve inpatient mortality. Appl Clin Inform 2010;1(4):394-407 [FREE Full text] [doi: 10.4338/ACI-2010-09-RA-0054] [Medline: 23616849]
56. Biltoft J, Restrepo C, Volk S, Casterton T, McLeese E. Use of an electronic early warning alert to improve sepsis outcomes in a community hospital. In: Proceedings of the 42nd Critical Care Congress of the Society of Critical Care Medicine, SCCM 2013. 2012 Presented at: 42nd Critical Care Congress of the Society of Critical Care Medicine, SCCM 2013; January 19-23, 2013; San Juan Puerto Rico URL: https://journals.lww.com/ccmjournal/Abstract/2012/12001/483__USE_OF_AN_ELECTRONIC_EARLY_WARNING_ALERT_TO.448.aspx [doi: 10.1097/01.ccm.0000424701.89139.e0]
57. Bradley P, DePolo M, Tipton J, Stacks H, Holt A. Improving sepsis survival using mews for early recognition and immediate response to patient decline. In: Proceedings of the 48th Critical Care Congress of the Society of Critical Care Medicine, SCCM 2019. 2019 Presented at: 48th Critical Care Congress of the Society of Critical Care Medicine, SCCM 2019; February 17-20, 2019; San Diego, CA, United States p. 788 URL: https://journals.lww.com/ccmjournal/Citation/2019/01001/1626__IMPROVING_SEPSIS_SURVIVAL_USING_MEWS_FOR.1578.aspx [doi: 10.1097/01.ccm.0000552367.92747.7a]
58. Brandt BN, Gartner AB, Moncure M, Cannon CM, Carlton E, Cleek C, et al. Identifying severe sepsis via electronic surveillance. Am J Med Qual 2015;30(6):559-565. [doi: 10.1177/1062860614541291] [Medline: 24970280]
59. Brown SM, Jones J, Kuttler KG, Keddington RK, Allen TL, Haug P. Prospective evaluation of an automated method to identify patients with severe sepsis or septic shock in the emergency department. BMC Emerg Med 2016 Aug 22;16(1):31 [FREE Full text] [doi: 10.1186/s12873-016-0095-0] [Medline: 27549755]
60. Buck KM. Developing an early sepsis alert program. J Nurs Care Qual 2014;29(2):124-132. [doi: 10.1097/NCQ.0b013e3182a98182] [Medline: 24048073]
61. Carlbom D, Kelly M. SePSIS: An innovative electronic warning system for in-hospital screening of SePSIS. In: Proceeings of the Critical Care Congress 2015. 2015 Presented at: Critical Care Congress 2015; January 17-21, 2015; Phoenix, AZ United States URL: https://journals.lww.com/ccmjournal/Fulltext/2014/12001/966__SEPSIS__AN_INNOVATIVE_ELECTRONIC_WARNING.933.aspx [doi: 10.1097/01.ccm.0000458463.39133.b9]
62. Chanas T, Volles D, Sawyer R, Mallow-Corbett S. Analysis of a new best-practice advisory on time to initiation of antibiotics in surgical intensive care unit patients with septic shock. J Intensive Care Soc 2019 Feb;20(1):34-39 [FREE Full text] [doi: 10.1177/1751143718767059] [Medline: 30792760]
63. Chang J, Sullivan M, Shea E, Shimabukuro D. The effectiveness of a real-time electronic alert to detect severe sepsis in an intensive care unit. In: Proceedings of the Annual Meeting of the International Anesthesia Research Society. 2015 Presented at: 2015 Annual Meeting of the International Anesthesia Research Society; March 21-24, 2015; Honolulu, HI United States p. S416. [doi: 10.1213/01.ane.0000470325.07465.0f]
64. Colorafi KJ, Ferrell K, D'Andrea A, Colorafi J. Influencing outcomes with automated time zero for sepsis through statistical validation and process improvement. Mhealth 2019;5:36 [FREE Full text] [doi: 10.21037/mhealth.2019.09.04] [Medline: 31620463]
65. Comlekoglu T, Nawzadi J, Glass G, Hartka T. Utility of electronic-and provider-initiated alerts in diagnosis of sepsis in the emergency department. In: Proceedings of the Annual Meeting of the Society for Academic Emergency Medicine,

SAEM 2020. 2020 Presented at: Annual Meeting of the Society for Academic Emergency Medicine, SAEM 2020; May 12-15, 2020; Denver, United States p. S210. [doi: 10.1111/acem.13961]
66. Croft CA, Moore FA, Efron PA, Marker PS, Gabrielli A, Westhoff LS, et al. Computer versus paper system for recognition and management of sepsis in surgical intensive care. J Trauma Acute Care Surg 2014 Feb;76(2):311-318. [doi: 10.1097/TA.0000000000000121] [Medline: 24458039]
67. Danak V. To improve sensitivity and specificity in early detection of sepsis in patients [Dissertation]. University of California, Davis. 2014. URL: https://www.proquest.com/docview/1665309377/abstract/403F4AE572954DA6PQ/ [accessed 2020-09-18]
68. Downing NL, Rolnick J, Poole SF, Hall E, Wessels AJ, Heidenreich P, et al. Electronic health record-based clinical decision support alert for severe sepsis: a randomised evaluation. BMJ Qual Saf 2019 Sep;28(9):762-768 [FREE Full text] [doi: 10.1136/bmjqs-2018-008765] [Medline: 30872387]
69. Dumont L, Harding AD. Development and implementation of a sepsis program. J Emerg Nurs 2013 Nov;39(6):625-630. [doi: 10.1016/j.jen.2013.08.009] [Medline: 24094734]
70. Dziadzko MA, Harrison AM, Tiong IC, Pickering BW, Moreno Franco P, Herasevich V. Testing modes of computerized sepsis alert notification delivery systems. BMC Med Inform Decis Mak 2016 Dec 09;16(1):156 [FREE Full text] [doi: 10.1186/s12911-016-0396-y] [Medline: 27938401]
71. Eck T, Kropf F, Rahmlow K, Pavlovic D, Wegner A, Guderian L, et al. The Sepsis-Tool helps to recognize severe sepsis earlier. In: Proceedings of the 4th International Congress of the German Sepsis Society "Sepsis and Multiorgan Dysfunction". 2009 Presented at: 4th International Congress of the German Sepsis Society "Sepsis and Multiorgan Dysfunction"; September 9-12, 2009; Weimer, Germany p. 65. [doi: 10.1007/s15010-009-1004-1]
72. Ehrlichman P, Trach L, Patel B, Maheshwari V, Seckel M. Sensitivity and positive predictive value of cerner EMR based sepsis recognition tool. In: Proceedings of the Critical Care Congress 2015. 2015 Presented at: Critical Care Congress 2015; January 17-21, 2015; Phoenix, AZ United States p. A1597 URL: https://journals.lww.com/ccmjournal/Fulltext/2014/12001/983__SENSITIVITY_AND_POSITIVE_PREDICTIVE_VALUE_OF.950.aspx [doi: 10.1097/01.ccm.0000458480.24778.4d]
73. Engineer R, Fertel B, Podolsky S, Gullett T, Smith E. Use of clinical decision support tools to meet compliance with CMS SEP-1 sepsis guidelines. In: Proceedings of the American College of Emergency Physicians, ACEP 2016 Research Forum. 2016 Oct Presented at: American College of Emergency Physicians, ACEP 2016 Research Forum; October 16-18, 2016; Las Vegas, NV United States p. S60. [doi: 10.1016/j.annemergmed.2016.08.161]
74. Eren G, Tekdos Y, Dogan M, Acicbe O, Kaya E, Hergunsel O. Septic shock alert over SIRS criteria has an impact on outcome but needs to be revised. In: Proceedings of the 36th International Symposium on Intensive Care and Emergency Medicine. 2016 Presented at: 36th International Symposium on Intensive Care and Emergency Medicine; March 15-18, 2016; Brussels, Belgium URL: https://cyberleninka.org/article/n/861694
75. Faisal M, Richardson D, Scally AJ, Howes R, Beatson K, Speed K, et al. Computer-aided National Early Warning Score to predict the risk of sepsis following emergency medical admission to hospital: a model development and external validation study. Can Med Assoc J 2019 Apr 08;191(14):382-389 [FREE Full text] [doi: 10.1503/cmaj.181418] [Medline: 30962196]
76. Falimirski M, Crews H, Broach T, Bosslet G. Early sepsis detection with an electronic medical record SIRS alert program. In: Proceedings of the 36th Annual Meeting of the Surgical Infection Society. 2016 Presented at: 36th Annual Meeting of the Surgical Infection Society; May 18-21, 2016; Palm Beach, Florida p. S30. [doi: 10.1089/sur.2016.29005.abstracts]
77. Fee C, Suess E, Maruoka A, Quon T, Meghan P, Maloney S. Effect of automated, real-time, electronic health record sirs and severe sepsis alerts on bundle compliance and mortality. In: Proceedings of the Annual Meeting of the Society for Academic Emergency Medicine, SAEM 2014. 2014 Presented at: Annual Meeting of the Society for Academic Emergency Medicine, SAEM 2014; May 13-17, 2014; Dallas, TX United States p. S102. [doi: 10.1111/acem.12365]
78. Fogerty R, Sankey C, Kenyon K, Sussman S, Acker K, Sigurdsson S, et al. Pilot of a low-resource, EHR-based protocol for sepsis monitoring, alert, and intervention. In: Proceedings of the 39th Annual Meeting of the Society of General Internal Medicine, SGIM 2016. 2016 Presented at: 39th Annual Meeting of the Society of General Internal Medicine, SGIM 2016; May 11-14, 2016; Hollywood, FL United States URL: https://shmabstracts.org/abstract/pilot-of-a-low-resource-ehr-based-protocol-for-sepsis-monitoring-alert-and-intervention/
79. Fogerty RL, Sussman LS, Kenyon K, Li F, Sukumar N, Kliger AS, et al. Using system inflammatory response syndrome as an easy-to-implement, sustainable, and automated tool for all-cause deterioration among medical inpatients. J Patient Saf 2019 Dec;15(4):74-77. [doi: 10.1097/PTS.0000000000000463] [Medline: 29369071]
80. Gatewood MO, Wemple M, Greco S, Kritek PA, Durvasula R. A quality improvement project to improve early sepsis care in the emergency department. BMJ Qual Saf 2015 Dec;24(12):787-795. [doi: 10.1136/bmjqs-2014-003552] [Medline: 26251506]
81. Gerald J, Alsip J, Hicks J, Waldrum M, Dunlap N. Using the EMR to perform continuous, automated, real-time surveillance to identify hospitalized patients at risk of sepsis. In: Poster Presentation : CHEST 2011. 2011 Oct Presented at: CHEST 2011; October 22-26, 2011; Honolulu, HI United States p. 426A. [doi: 10.1378/chest.1118563]
82. Gerald J, Alsip J, Hicks J, Waldrum M, Dunlap N. Predictors of in-hospital mortality among patients identified as having severe sepsis by an automated clinical decision rule. In: Proceedings of the American Thoracic Society International Conference, ATS 2011. 2011 Presented at: American Thoracic Society International Conference, ATS 2011; May 13-18,

2011; Denver, CO, United States URL: https://www.atsjournals.org/doi/abs/10.1164/ajrccm-conference.2011.183.1_MeetingAbstracts.A4696 [doi: 10.1164/ajrccm-conference.2011.183.1_meetingabstracts.a4696]
83. Gerald J, Alsip J, Hicks J, Waldrum M, Dunlap N. Pattern of clinical and laboratory abnormalities among hospitalized patients identified as having severe sepsis by an automated clinical decision rule. In: Proceedings of the American Thoracic Society International Conference, ATS 2011. 2011 Presented at: American Thoracic Society International Conference, ATS 2011; May 13-18, 2011; Denver, CO, United States URL: https://www.atsjournals.org/doi/abs/10.1164/ajrccm-conference.2011.183.1_MeetingAbstracts.A3860 [doi: 10.1164/ajrccm-conference.2011.183.1_meetingabstracts.a3860]
84. Ghanem-Zoubi NO, Vardi M, Laor A, Weber G, Bitterman H. Assessment of disease-severity scoring systems for patients with sepsis in general internal medicine departments. Crit Care 2011;15(2):R95 [FREE Full text] [doi: 10.1186/cc10102] [Medline: 21401927]
85. Giesler E, Sherwin R, Courage C, Stewart S, Fiorvento A, Powell A, et al. Use of computerized sepsis care plans improved resuscitation parameters in patients with severe sepsis and septic shock. In: Proceedings of the American College of Emergency Physicians, ACEP 2014 Research Forum. 2014 Presented at: American College of Emergency Physicians, ACEP 2014 Research Forum; October 27-28, 2014; Chicago, IL, United States p. S41. [doi: 10.1016/j.annemergmed.2014.07.139]
86. Gille J, Dietz A, Taha H, Sablotzki A. A sirs-based automated alarm system for the diagnosis of sepsis after burn injury. Ann Burns Fire Disasters 2017 Sep 30;30(3):177-184 [FREE Full text] [Medline: 29849519]
87. Gowda SS, Hooper M, Weavind L, Nian H, Wheeler A. Modified systemic inflammatory response syndrome (MSIRS) score and risk of mortality in medical intensive care unit (MICU) patients. In: Proceedings of the Critical Care Congress 2012. 2012 Presented at: Critical Care Congress 2012; February 4-8, 2012; Houston, TX, United States p. 154 URL: https://journals.lww.com/ccmjournal/Citation/2011/12001/556__MODIFIED_SYSTEMIC_INFLAMMATORY_RESPONSE.514.aspx
88. Gowda SS, Nian H, Weavind L. Modified SIRS criteria in peri-operative patients admitted to surgical intensive care unit: clinical course and outcome. In: Proceedings of the Critical Care Congress 2012. 2012 Presented at: Critical Care Congress 2012; February 4-8, 2012; Houston, TX, United States p. 154 URL: https://journals.lww.com/ccmjournal/Citation/2011/12001/557__MODIFIED_SIRS_CRITERIA_IN_PERI_OPERATIVE.515.aspx
89. Grek A, Booth S, Festic E, Maniaci M, Shirazi E, Thompson K, et al. Sepsis and shock response team: impact of a multidisciplinary approach to implementing surviving sepsis campaign guidelines and surviving the process. Am J Med Qual 2017;32(5):500-507. [doi: 10.1177/1062860616676887] [Medline: 27837163]
90. Guidi JL, Clark K, Upton MT, Faust H, Umscheid CA, Lane-Fall MB, et al. Clinician perception of the effectiveness of an automated early warning and response system for sepsis in an academic medical center. Ann Am Thorac Soc 2015 Oct;12(10):1514-1519. [doi: 10.1513/AnnalsATS.201503-129OC] [Medline: 26288388]
91. Guirgis FW, Jones L, Esma R, Weiss A, McCurdy K, Ferreira J, et al. Managing sepsis: electronic recognition, rapid response teams, and standardized care save lives. J Crit Care 2017 Aug;40:296-302 [FREE Full text] [doi: 10.1016/j.jcrc.2017.04.005] [Medline: 28412015]
92. Harrison AM. Development, testing, and refining the failure to rescue sepsis sniffer [Dissertation]. College of Medicine, Mayo Clinic. 2015. URL: https://www.proquest.com/openview/d6cf865a3f8eb3c4ed3583317f9ff53b/1?pq-origsite=gscholar&cbl=18750 [accessed 2020-09-18]
93. Harrison AM, Thongprayoon C, Aakre CA, Jeng JY, Dziadzko MA, Gajic O, et al. Comparison of methods of alert acknowledgement by critical care clinicians in the ICU setting. PeerJ 2017;5:e3083 [FREE Full text] [doi: 10.7717/peerj.3083] [Medline: 28316887]
94. Hayden GE, Tuuri RE, Scott R, Losek JD, Blackshaw AM, Schoenling AJ, et al. Triage sepsis alert and sepsis protocol lower times to fluids and antibiotics in the ED. Am J Emerg Med 2016 Jan;34(1):1-9 [FREE Full text] [doi: 10.1016/j.ajem.2015.08.039] [Medline: 26386734]
95. Herasevich V, Pieper MS, Pulido J, Gajic O. Enrollment into a time sensitive clinical study in the critical care setting: results from computerized septic shock sniffer implementation. J Am Med Inform Assoc 2011;18(5):639-644 [FREE Full text] [doi: 10.1136/amiajnl-2011-000228] [Medline: 21508415]
96. Hiensch R, Poeran J, Saunders-Hao P, Adams V, Powell CA, Glasser A, et al. Impact of an electronic sepsis initiative on antibiotic use and health care facility-onset Clostridium difficile infection rates. Am J Infect Control 2017 Oct 01;45(10):1091-1100. [doi: 10.1016/j.ajic.2017.04.005] [Medline: 28602274]
97. Honeyford K, Cooke GS, Kinderlerer A, Williamson E, Gilchrist M, Holmes A, Sepsis Big Room, et al. Evaluating a digital sepsis alert in a London multisite hospital network: a natural experiment using electronic health record data. J Am Med Inform Assoc 2020 Feb 01;27(2):274-283 [FREE Full text] [doi: 10.1093/jamia/ocz186] [Medline: 31743934]
98. Hooper MH, Martin JB, Weavind LM, Semler MW, Albert DW, Deane NB, et al. Automated surveillance of modified sirs criteria is an effective tool for detection of sepsis in the medical intensive care unit. In: Proceedings of the American Thoracic Society International Conference, ATS 2010. 2010 Presented at: American Thoracic Society International Conference, ATS 2010; May 14-19, 2010; New Orleans, LA, United States. [doi: 10.1164/ajrccm-conference.2010.181.1_meetingabstracts.a6137]

99. Hooper MH. Modified sirs criteria for detection of sepsis in the intensive care unit. In: Proceedings of the American Thoracic Society International Conference, ATS 2011. 2011 Presented at: American Thoracic Society International Conference, ATS 2011; May 13-18, 2011; Denver, CO, United States. [doi: 10.1164/ajrccm-conference.2011.183.1_meetingabstracts.a5848]
100. Hooper MH, Weavind L, Wheeler AP, Martin JB, Gowda SS, Semler MW, et al. Randomized trial of automated, electronic monitoring to facilitate early detection of sepsis in the intensive care unit*. Crit Care Med 2012 Jul;40(7):2096-2101 [FREE Full text] [doi: 10.1097/CCM.0b013e318250a887] [Medline: 22584763]
101. Horton DJ, Graves KK, Kukhareva PV, Johnson SA, Cedillo M, Sanford M, et al. Modified early warning score-based clinical decision support: cost impact and clinical outcomes in sepsis. JAMIA Open 2020 Jul;3(2):261-268 [FREE Full text] [doi: 10.1093/jamiaopen/ooaa014] [Medline: 32734167]
102. Huff S, Stephens K, Whiteman K, Swanson-Biearman B, Mori C. Implementation of a vital sign alert system to improve outcomes. J Nurs Care Qual 2019;34(4):346-351. [doi: 10.1097/NCQ.0000000000000384] [Medline: 30585982]
103. Hughes MC, Roedocker A, Ehli J, Walz D, Froehlich K, White L, et al. A quality improvement project to improve sepsis-related outcomes at an integrated healthcare system. J Healthc Qual 2019;41(6):369-375. [doi: 10.1097/JHQ.0000000000000193] [Medline: 30883463]
104. Jones SL, Ashton CM, Kiehne L, Gigliotti E, Bell-Gordon C, Disbot M, et al. Reductions in sepsis mortality and costs after design and implementation of a nurse-based early recognition and response program. Jt Comm J Qual Patient Saf 2015 Nov;41(11):483-491 [FREE Full text] [doi: 10.1016/s1553-7250(15)41063-3] [Medline: 26484679]
105. Judd WR, Stephens DM, Kennedy CA. Clinical and economic impact of a quality improvement initiative to enhance early recognition and treatment of sepsis. Ann Pharmacother 2014 Oct;48(10):1269-1275. [doi: 10.1177/1060028014541792] [Medline: 24982314]
106. Jukes T. Using clinical decision support to assist in identification of severe sepsis: design development and implementation of an automated sepsis prompt [Dissertation]. University of California, Davis. 2013. URL: https://www.proquest.com/docview/1449817718? [accessed 2020-09-18]
107. Jung AD, Baker J, Droege CA, Nomellini V, Johannigman J, Holcomb JB, et al. Sooner is better: use of a real-time automated bedside dashboard improves sepsis care. J Surg Res 2018 Nov;231:373-379. [doi: 10.1016/j.jss.2018.05.078] [Medline: 30278956]
108. Kangas C, Iverson L, Pierce D. Sepsis screening: combining early warning scores and SIRS criteria. Clin Nurs Res 2021 Jan;30(1):42-49. [doi: 10.1177/1054773818823334] [Medline: 30654646]
109. Krieger P, Jackson J, Sarin R, Wang C, Acquah S, Smith M, et al. A novel electronic medical record-based screening alert successfully identifies patients with severe sepsis. In: Proceedings of the American College of Emergency Physicians, ACEP Research Forum 2013. 2013 Oct Presented at: American College of Emergency Physicians, ACEP Research Forum 2013; October 14-15, 2013; Seattle, WA, United States p. S96. [doi: 10.1016/j.annemergmed.2013.07.087]
110. Kurczewski L, Sweet M, McKnight R, Halbritter K. Reduction in time to first action as a result of electronic alerts for early sepsis recognition. Crit Care Nurs Q 2015;38(2):182-187. [doi: 10.1097/CNQ.0000000000000060] [Medline: 25741959]
111. Lawson L. Use of clinical decision algorithms on medical surgical units to assist in the identification and care of patients at risk for acquiring sepsis [Dissertation]. Northcentral University. 2017. URL: https://www.proquest.com/openview/9de3653d06b2e4128ccde607b239edf2/1?pq-origsite=gscholar&cbl=18750&diss=y [accessed 2020-09-18]
112. Liang D, Cope J, Patel R, Holder M. Comparison of two sepsis alert systems within an academic medical center. In: Proceedings of the 47th Society of Critical Care Medicine Critical Care Congress, SCCM 2018. 2018 Presented at: 47th Society of Critical Care Medicine Critical Care Congress, SCCM 2018; February 25-28, 2018; San Antonio, TX, United States p. 472. [doi: 10.1097/01.ccm.0000528984.10458.e7]
113. MacMillan A, Rudinsky D, Han G, Elliott JO, Jordan K. Multidisciplinary approach to improve sepsis outcomes. J Healthc Qual 2019;41(4):220-227. [doi: 10.1097/JHQ.0000000000000166] [Medline: 30418300]
114. Mann-Salinas E, Caldwell N, Serio-Melvin M, Luellen D, Chung K, Cancio L, et al. Evaluation of burn sepsis automated alerts in an intensive care unit. In: Proceedings of the Critical Care Congress 2015. 2015 Presented at: Critical Care Congress 2015; January 17-21, 2015; Phoenix, AZ, United States p. A1596. [doi: 10.1097/01.ccm.0000458478.79036.7c]
115. Manteuffel J, Coba V, Jaehne K. An emergency department sepsis alert is associated with decreased mortality in septic patients admitted to the intensive care unit. In: Proceedings of the American College of Emergency Physicians, ACEP Research Forum 2013. 2013 Oct Presented at: American College of Emergency Physicians, ACEP Research Forum 2013; October 14-15, 2013; Seattle, WA, United States p. S63. [doi: 10.1016/j.annemergmed.2013.07.456]
116. Martin-Rico P, Valdivia-Perez A, Lacalle-Martinez J, Chorda-Ribelles J, Marco-Lattur M, Lozano-Cortell N, et al. Assessing the value of a real-time electronic screening algorithm for early detection of severe sepsis in the emergency department. Crit Care 2014 Dec 3;18(S2):P22. [doi: 10.1186/cc14025]
117. McCabe P, Team R, Thorne J, Sarumi M, Ryan A, Rappaport H. M2SEWS early warning systems and proactive screening provides early intervention and escalates care. In: Proceedings of the 45th Critical Care Congress of the Society of Critical Care Medicine, SCCM 2015. 2015 Presented at: 45th Critical Care Congress of the Society of Critical Care Medicine, SCCM 2015; February 20-24, 2016; Orlando, FL, United States p. 276-277 URL: https://journals.lww.com/ccmjournal/Citation/2015/12001/1100__M2SEWS_EARLY_WARNING_SYSTEMS_AND_PROACTIVE.1103.aspx

118. McGrane T, Rice T, Hooper M, Gowda S, Weavind L. Electronic SIRS alert facilitating recognition of sepsis by housestaff. In: Proceedings of the 40th Critical Care Congress of the Society of Critical Care Medicine. 2010 Presented at: 40th Critical Care Congress of the Society of Critical Care Medicine; January 15-19, 2011; San Diego, CA, United States p. A170. [doi: 10.1097/01.ccm.0000390903.16849.8c]
119. McRee L, Thanavaro JL, Moore K, Goldsmith M, Pasvogel A. The impact of an electronic medical record surveillance program on outcomes for patients with sepsis. Heart Lung 2014;43(6):546-549. [doi: 10.1016/j.hrtlng.2014.05.009] [Medline: 24961633]
120. Meurer WJ, Smith BL, Losman ED, Sherman D, Yaksich JD, Jared JD, et al. Real-time identification of serious infection in geriatric patients using clinical information system surveillance. J Am Geriatr Soc 2009 Jan;57(1):40-45. [doi: 10.1111/j.1532-5415.2008.02094.x] [Medline: 19170788]
121. Miller K, Kowalski R, Capan M, Wu P, Mosby D, Arnold R. Assessment of nursing response to a real-time alerting tool for sepsis: a provider survey. Am J Hosp Med 2017;1(3) [FREE Full text] [Medline: 30854401]
122. Muratore S, Coughlan A, Glover JK, Statz CL, Weinert C, Kline SE, et al. Electronically triggered sepsis alert in non-intensive care inpatients using modified systemic inflammatory response syndrome criteria: a retrospective observational study with in-depth analysis of surgical patients. Surg Infect (Larchmt) 2019;20(4):278-285. [doi: 10.1089/sur.2018.228] [Medline: 30724713]
123. Narayanan N, Gross AK, Pintens M, Fee C, MacDougall C. Effect of an electronic medical record alert for severe sepsis among ED patients. Am J Emerg Med 2016 Feb;34(2):185-188. [doi: 10.1016/j.ajem.2015.10.005] [Medline: 26573784]
124. Nelson JL, Smith BL, Jared JD, Younger JG. Prospective trial of real-time electronic surveillance to expedite early care of severe sepsis. Ann Emerg Med 2011 May;57(5):500-504. [doi: 10.1016/j.annemergmed.2010.12.008] [Medline: 21227543]
125. Nelson G, Kamash T, Sherwin R. The performance of an electronic medical record-based alert to identify emergency department patients with severe sepsis and septic shock. In: Proceedings of the American College of Emergency Physicians, ACEP Research Forum 2012. 2012 Oct Presented at: American College of Emergency Physicians, ACEP Research Forum 2012; October 8-9, 2012; Denver, CO, United States p. S21. [doi: 10.1016/j.annemergmed.2012.06.332]
126. Nguyen SQ, Mwakalindile E, Booth JS, Hogan V, Morgan J, Prickett CT, et al. Automated electronic medical record sepsis detection in the emergency department. PeerJ 2014;2:e343 [FREE Full text] [doi: 10.7717/peerj.343] [Medline: 24765577]
127. Noonan TZ. Decision making under uncertainty in the emergency department: studying the effects of cognitive biases in the diagnosis of sepsis [Dissertation]. The University of Iowa. 2018. URL: https://www.proquest.com/openview/56541ec710f706fd3432237557287af3/1?pq-origsite=gscholar&cbl=18750 [accessed 2020-09-18]
128. Okafor N, Mehta A, Freeman C, Doshi P. Evaluation of a sirs alert tool as a method for early detection of sepsis in the emergency department. In: Proceedings of the American College of Emergency Physicians, ACEP 2014 Research Forum. 2014 Presented at: American College of Emergency Physicians, ACEP 2014 Research Forum; October 27-28, 2014; Chicago, IL, United States p. S56. [doi: 10.1016/j.annemergmed.2014.07.180]
129. Oxman D, Oettinger G, Pugliese R, El Beyrouty C, McDaniel C, Riggio J, et al. An interdisciplinary program for improving the recognition and treatment of severe sepsis. In: Proceedings of the 43rd Critical Care Congress of the Society of Critical Care Medicine, SCCM 2014. 2013 Presented at: 43rd Critical Care Congress of the Society of Critical Care Medicine, SCCM 2014; January 9-13, 2014; San Francisco, CA, United States p. A270. [doi: 10.1097/01.ccm.0000440308.10021.b8]
130. Pandit A, Hale K, Pottinger T, Bugov D, Lacaille S, Rebal S, et al. Analysis of an automated sepsis screening alert at a tertiary cancer center. In: Poster Presentation: CHEST 2015. 2015 Oct Presented at: CHEST 2015; October 24-28, 2015; Montreal, QC, Canada p. 342A. [doi: 10.1378/chest.2274431]
131. Perlin JB, Jackson E, Hall C, Mindick A, Berberich G, Korwek K, et al. 2019 John M. Eisenberg Patient Safety and Quality Awards: SPOTting sepsis to save lives: a nationwide computer algorithm for early detection of sepsis: innovation in patient safety and quality at the national level (Eisenberg Award). Jt Comm J Qual Patient Saf 2020 Jul;46(7):381-391. [doi: 10.1016/j.jcjq.2020.04.006] [Medline: 32598281]
132. Pertiwi AA, Fraczkowski D, Stogis SL, Lopez KD. Using heuristic evaluation to improve sepsis alert usability. Crit Care Nurs Clin North Am 2018 Jun;30(2):297-309. [doi: 10.1016/j.cnc.2018.02.011] [Medline: 29724447]
133. Pham T, Kupchik N, Rivera S, Robinson C, Shea N. EB44 automated sepsis screening in a progressive care unit. Critical Care Nurse 2012;32(2):27-28 [FREE Full text]
134. Powers B, Mwangi J, Ordaz E, Weber L, patel N, Uppalapati A, et al. EMR-Based sepsis alert: improving early detection. In: Proceedings of the Critical Care Congress 2015. 2015 Presented at: Critical Care Congress 2015; January 17-21, 2015; Phoenix, AZ, United States p. 1598-1599. [doi: 10.1097/01.ccm.0000458488.62896.96]
135. Pulia M, Vuong I, Schandel C, Sharp B. Impact of an electronic health record sepsis screen on antibiotic stewardship in the emergency department. In: Proceedings of the Annual Meeting of the Society for Academic Emergency Medicine, SAEM 2016. 2016 Presented at: Annual Meeting of the Society for Academic Emergency Medicine, SAEM 2016; May 10-13, 2016; New Orleans, LA, United States p. S147. [doi: 10.1111/acem.12974]
136. Rauch S, Fischer M, Martin C, Sablotzki A. Automated, electronic monitoring for early detection of sepsis. In: Proceedings of the 33rd International Symposium on Intensive Care and Emergency Medicine. 2013 Presented at: 33rd International Symposium on Intensive Care and Emergency Medicine; March 19-22, 2013; Brussels, Belgium p. S191. [doi: 10.1186/cc12457]

137. Rincon TA, Manos EL, Pierce JD. Telehealth intensive care unit nurse surveillance of sepsis. Comput Inform Nurs 2017 Sep;35(9):459-464. [doi: 10.1097/CIN.0000000000000355] [Medline: 28445172]
138. Salamah M, Thagafi M, Reidy M, Diknash S, Al-Qahtani S, Marini A, et al. The predictive value of an electronic alert system for severe sepsis and septic shock. In: Proceedings of the American Thoracic Society International Conference, ATS 2013. 2013 Presented at: American Thoracic Society International Conference, ATS 2013; May 17-22, 2013; Philadelphia, PA, United States URL: https://www.atsjournals.org/doi/abs/10.1164/ajrccm-conference.2013.187.1_MeetingAbstracts.A3081
139. Sankey C, Sussman LS, Kenyon K, Li F, Sukumar N, Kliger A, et al. Pilot of a low-resource, EHR-based tool for sepsis monitoring, alert, and intervention. In: Proceedings of the 40th Annual Meeting of the Society of General Internal Medicine, SGIM 2017. 2017 Presented at: 40th Annual Meeting of the Society of General Internal Medicine, SGIM 2017; April 19-22, 2017; Washington, DC, United States URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5391321/
140. Santistevan J, Hamedani A, Werner E, Spranger J, Sharp B. Performance of an automated severe sepsis screening tool in the emergency department. In: Proceedings of the American College of Emergency Physicians, ACEP 2017 Research Forum. 2017 Presented at: American College of Emergency Physicians, ACEP 2017 Research Forum; October 29-31, 2017; Washington, DC, United States p. 163-164. [doi: 10.1016/j.annemergmed.2017.07.296]
141. Sarumi M, Garner B, Rappaport H, McCabe P, Gutierrez S, Ryan A, et al. Implementation of a standardized and reliable system improves mortality in patients with severe sepsis. In: Proceedings of the 43rd Critical Care Congress of the Society of Critical Care Medicine, SCCM 2014. 2014 Presented at: 43rd Critical Care Congress of the Society of Critical Care Medicine, SCCM 2014; January 9-13, 2014; San Francisco, CA, United States p. A273. [doi: 10.1097/01.ccm.0000440317.37837.1c]
142. Sawyer AM, Deal EN, Labelle AJ, Witt C, Thiel SW, Heard K, et al. Implementation of a real-time computerized sepsis alert in nonintensive care unit patients. Crit Care Med 2011 Mar;39(3):469-473. [doi: 10.1097/CCM.0b013e318205df85] [Medline: 21169824]
143. Seetharaman S, Wilson C, Landrum M, Qasba S, Katz M, Ladikos N, et al. Does use of electronic alerts for systemic inflammatory response syndrome (sirs) to identify patients with sepsis improve mortality? Am J Med 2019 Jul;132(7):862-868. [doi: 10.1016/j.amjmed.2019.01.032] [Medline: 30831065]
144. Sharp B, Schandel C, Vuong I, Masters M, Lee A, Pothof J, et al. Implementation of an electronic ED sepsis screen and alert: effect on compliance with ED sepsis quality measures. In: Proceedings of the Annual Meeting of the Society for Academic Emergency Medicine, SAEM 2016. 2016 Presented at: Annual Meeting of the Society for Academic Emergency Medicine, SAEM 2016; May 10-13, 2016; New Orleans, LA, United States p. S96. [doi: 10.1111/acem.12974]
145. Sharp B, Schandel C, Vuong I, Masters M, Lee A, Pothof J, et al. Silent no longer: sepsis recognition by electronic screening in the emergency department. In: Proceedings of the Annual Meeting of the Society for Academic Emergency Medicine, SAEM 2016. 2016 Presented at: Annual Meeting of the Society for Academic Emergency Medicine, SAEM 2016; May 10-13, 2016; New Orleans, LA, United States p. 95-96. [doi: 10.1111/acem.12974]
146. Sherwin R, Courage C, Stewart S, Robinson D, Ying H. Performance of two clinical decision support tools to identify sepsis patients in the emergency department. In: Proceedings of the Annual Meeting of the Society for Academic Emergency Medicine, SAEM 2014. 2014 Presented at: Annual Meeting of the Society for Academic Emergency Medicine, SAEM 2014; May 13-17, 2014; Dallas, TX, United States p. S279. [doi: 10.1111/acem.12365]
147. Shetty A, Murphy M, Middleton-Rennie C, Lancuba A, Green M, Lander H, et al. Evaluation of an augmented emergency department electronic medical record-based sepsis alert. Emerg Med Australas 2021 Feb 23. [doi: 10.1111/1742-6723.13748] [Medline: 33622028]
148. Song J, Cho H, Park DW, Ahn S, Kim JY, Seok H, et al. The effect of the intelligent sepsis management system on outcomes among patients with sepsis and septic shock diagnosed according to the sepsis-3 definition in the emergency department. J Clin Med 2019 Oct 27;8(11):27 [FREE Full text] [doi: 10.3390/jcm8111800] [Medline: 31717855]
149. Swenson K, Ferguson J, Shieh L, Rogers A. Attitudes and perceptions of medical trainees towards an electronic medical alert system for sepsis. In: Proceedings of the American Thoracic Society International Conference, ATS 2018. 2018 Presented at: American Thoracic Society International Conference, ATS 2018; May 18-23, 2018; San Diego, CA, United States URL: https://www.atsjournals.org/doi/abs/10.1164/ajrccm-conference.2018.197.1_MeetingAbstracts.A4504
150. Swenson K, Rogers A, Krishnan G, Shieh L. Analysis of a potential automated sepsis alert based on quick sequential organ failure assessment (qSOFA) among hospitalized floor patients. In: Proceedings of the American Thoracic Society 2019 International Conference. 2019 Presented at: American Thoracic Society 2019 International Conference; May 17-22, 2019; Dallas, TX, United States. [doi: 10.1164/ajrccm-conference.2019.199.1_MeetingAbstracts.A7173]
151. Tafelski S, Nachtigall I, Deja M, Tamarkin A, Trefzer T, Halle E, et al. Computer-assisted decision support for changing practice in severe sepsis and septic shock. J Int Med Res 2010;38(5):1605-1616 [FREE Full text] [doi: 10.1177/147323001003800505] [Medline: 21309474]
152. Thompson DS, Oberteuffer R, Dorman T. Sepsis alert and diagnostic system: integrating clinical systems to enhance study coordinator efficiency. Comput Inform Nurs 2003;21(1):22-27. [doi: 10.1097/00024665-200301000-00009] [Medline: 12544151]

153. Tsai C, Patel K, Vincent A, Verzosa N, Norris D, Tillis W, et al. Electronic best practice advisories' effectiveness in detecting sepsis in the emergency department. Ann Emerg Med 2015 Oct;66(4):91-92. [doi: 10.1016/j.annemergmed.2015.07.287]
154. Umscheid CA, Betesh J, VanZandbergen C, Hanish A, Tait G, Mikkelsen ME, et al. Development, implementation, and impact of an automated early warning and response system for sepsis. J Hosp Med 2015 Jan;10(1):26-31 [FREE Full text] [Medline: 25263548]
155. Watanabe M, Kim B, Jeng A. St John sepsis alert correlation with management of infections and antimicrobial selection: a pilot study. In: Proceedings of the Open Forum Infectious Diseases, ID Week 2017. 2017 Presented at: Open Forum Infectious Diseases, ID Week 2017; October 4-8, 2017; San Diego, CA, United States p. S513. [doi: 10.1093/ofid/ofx163.1333]
156. Wawrose R, Baraniuk M, Standiford L, Wade C, Holcomb J, Moore L. Comparison of sepsis screening tools' ability to detect sepsis accurately. Surg Infect (Larchmt) 2016 Oct;17(5):525-529. [doi: 10.1089/sur.2015.069] [Medline: 27447053]
157. Westphal GA, Pereira AB, Fachin SM, Sperotto G, Gonçalves M, Albino L, et al. An electronic warning system helps reduce the time to diagnosis of sepsis. Rev Bras Ter Intensiva 2018;30(4):414-422 [FREE Full text] [Medline: 30570029]
158. Winterbottom F. Early warning: detecting sepsis in every patient, every shift requires timely vital sign documentation. In: Proceedings of the Ochsner 13th Annual Research Day Abstracts. 2016 Presented at: Ochsner 13th Annual Research Day Abstracts; May 17, 2016; New Orleans, LA, United States p. 398 URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5024823/
159. Wolpaw B, Kress N, Grant R, Wahl D, Garver-Hume A, Villanueva S, et al. EHR-based sepsis warning system improves detection and treatment of hospital-acquired sepsis in an acute care hospital setting. In: Proceedings of the 41st Annual Meeting of the Society of General Internal Medicine, SGIM 2018. 2018 Presented at: 41st Annual Meeting of the Society of General Internal Medicine, SGIM 2018; April 11-14, 2018; Denver, CO, United States URL: https://link.springer.com/content/pdf/10.1007%2Fs11606-018-4413-y.pdf
160. Yeich T, Brown S, Thomas K, Walker R. The sepsis alert: real-time electronic health record surveillance in the emergency department and sepsis outcomes. In: Proceedings of the Annual Meeting of the Society for Academic Emergency Medicine, SAEM 2012. 2012 Presented at: Annual Meeting of the Society for Academic Emergency Medicine, SAEM 2012; May 9-12, 2012; Chicago, IL, United States p. S49. [doi: 10.1111/j.1553-2712.2012.01332.x]
161. Yohannes S. Multimodal sepsis performance improvement initiative improves early and appropriate treatment, reduces sepsis-related readmissions, and reduces overall mortality. In: Proceedings of the Open Forum Infectious Diseases, ID Week 2019. 2019 Presented at: Open Forum Infectious Diseases, ID Week 2019; October 2-6, 2019; Washington, DC, United States p. S343. [doi: 10.1093/ofid/ofz360.838]
162. Zaragoza R, Sancho S, Hurtado C, Camarena J, González R, Borrás S, et al. SEPSIS SAVER. Preliminary results of a pilot implementation of a new automatic electronic alert for early detection of sepsis based on Sepsis 3 definition. In: Proceedings of the 32nd European Society of Intensive Care Medicine Annual Congress, ESICM 2019. 2019 Presented at: 32nd European Society of Intensive Care Medicine Annual Congress, ESICM 2019; September 28 - October 2, 2019; Berlin, Germany URL: https://icm-experimental.springeropen.com/articles/supplements/volume-7-supplement-3
163. Zuick S, Asch DA, Graustein A, Urbani R, Carrol B. Can a computerized sepsis screening and alert system accurately diagnose sepsis in hospitalized floor patients and potentially provide opportunities for early intervention? A pilot study. Intens Crit Care 2016;02(03):1-7. [doi: 10.21767/2471-8505.100049]
164. Uffen JW, Oosterheert JJ, Schweitzer VA, Thursky K, Kaasjager HA, Ekkelenkamp MB. Interventions for rapid recognition and treatment of sepsis in the emergency department: a narrative review. Clin Microbiol Infect 2020 Feb 29. [doi: 10.1016/j.cmi.2020.02.022] [Medline: 32120030]
165. Despins LA. Automated detection of sepsis using electronic medical record data: a systematic review. J Healthc Qual 2016 Sep 13. [doi: 10.1097/JHQ.0000000000000066] [Medline: 27631709]
166. Warttig S, Alderson P, Evans DJ, Lewis SR, Kourbeti IS, Smith AF. Automated monitoring compared to standard care for the early detection of sepsis in critically ill patients. Cochrane Database Syst Rev 2018 Jun 25;6:CD012404 [FREE Full text] [doi: 10.1002/14651858.CD012404.pub2] [Medline: 29938790]
167. Feldman SS, Buchalter S, Hayes LW. Health information technology in healthcare quality and patient safety: literature review. JMIR Med Inform 2018 Jun 04;6(2):e10264 [FREE Full text] [doi: 10.2196/10264] [Medline: 29866642]
168. Petticrew M, Knai C, Thomas J, Rehfuess EA, Noyes J, Gerhardus A, et al. Implications of a complexity perspective for systematic reviews and guideline development in health decision making. BMJ Glob Health 2019;4(Suppl 1):e000899 [FREE Full text] [doi: 10.1136/bmjgh-2018-000899] [Medline: 30775017]
169. Plsek PE, Greenhalgh T. Complexity science: the challenge of complexity in health care. Br Med J 2001 Sep 15;323(7313):625-628 [FREE Full text] [doi: 10.1136/bmj.323.7313.625] [Medline: 11557716]
170. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
171. Arefian H, Heublein S, Scherag A, Brunkhorst FM, Younis MZ, Moerer O, et al. Hospital-related cost of sepsis: a systematic review. J Infect 2017 Feb;74(2):107-117. [doi: 10.1016/j.jinf.2016.11.006] [Medline: 27884733]

172. Torio C, Andrews R. National inpatient hospital costs: the most expensive conditions by payer. In: Healthcare Cost and Utilization Project (HCUP) Statistical Briefs. Rockville (MD): Agency for Healthcare Research and Quality (US); 2013.
173. Peute LW, Spithoven R, Bakker PJ, Jaspers MW. Usability studies on interactive health information systems; where do we stand? Stud Health Technol Inform 2008;136:327-332. [Medline: 18487752]
174. Miller K, Mosby D, Capan M, Kowalski R, Ratwani R, Noaiseh Y, et al. Interface, information, interaction: a narrative review of design and functional requirements for clinical decision support. J Am Med Inform Assoc 2018 May 01;25(5):585-592 [FREE Full text] [doi: 10.1093/jamia/ocx118] [Medline: 29126196]
175. Hardenbol AX, Knols B, Louws M, Meulendijk M, Askari M. Usability aspects of medication-related decision support systems in the outpatient setting: a systematic literature review. Health Informatics J 2020 Mar;26(1):72-87 [FREE Full text] [doi: 10.1177/1460458218813732] [Medline: 30497336]
176. Harrison AM, Gajic O, Pickering BW, Herasevich V. Development and implementation of sepsis alert systems. Clin Chest Med 2016 Jun;37(2):219-229 [FREE Full text] [doi: 10.1016/j.ccm.2016.01.004] [Medline: 27229639]
177. The Lancet Respiratory Medicine. Crying wolf: the growing fatigue around sepsis alerts. Lancet Respir Med 2018 Mar;6(3):161. [doi: 10.1016/S2213-2600(18)30072-9] [Medline: 29508700]
178. Khairat S, Marc D, Crosby W, Al Sanousi A. Reasons for physicians not adopting clinical decision support systems: critical analysis. JMIR Med Inform 2018 Apr 18;6(2):e24 [FREE Full text] [doi: 10.2196/medinform.8912] [Medline: 29669706]
179. Jones D, Duke G, Green J, Briedis J, Bellomo R, Casamento A, et al. Medical emergency team syndromes and an approach to their management. Crit Care 2006 Feb;10(1):R30 [FREE Full text] [doi: 10.1186/cc4821] [Medline: 16507153]
180. Henning DJ, Oedorf K, Day DE, Redfield CS, Huguenel CJ, Roberts JC, et al. Derivation and validation of predictive factors for clinical deterioration after admission in emergency department patients presenting with abnormal vital signs without shock. West J Emerg Med 2015 Dec;16(7):1059-1066 [FREE Full text] [doi: 10.5811/westjem.2015.9.27348] [Medline: 26759655]
181. Hajj J, Blaine N, Salavaci J, Jacoby D. The "Centrality of Sepsis": a review on incidence, mortality, and cost of care. Healthcare (Basel) 2018 Jul 30;6(3) [FREE Full text] [doi: 10.3390/healthcare6030090] [Medline: 30061497]

Abbreviations

CCDS: computerized clinical decision support
ICU: intensive care unit
LILACS: Latin American and Caribbean Health Sciences Literature
PQDT: ProQuest Dissertations and Theses Global
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
SIRS: systemic inflammatory response syndrome

Edited by R Kukafka; submitted 17.06.21; peer-reviewed by K Honeyford, J Walsh; comments to author 30.07.21; revised version received 23.09.21; accepted 29.10.21; published 23.02.22.

Please cite as:
Ackermann K, Baker J, Green M, Fullick M, Varinli H, Westbrook J, Li L
Computerized Clinical Decision Support Systems for the Early Detection of Sepsis Among Adult Inpatients: Scoping Review
J Med Internet Res 2022;24(2):e31083
URL: https://www.jmir.org/2022/2/e31083
doi:10.2196/31083
PMID:35195528

Viewpoint

Understanding Decision-Making in the Adoption of Digital Health Technology: The Role of Behavioral Economics' Prospect Theory

Waqas Ullah Khan[1,2], BA, BEd, BSc, MSc, MD; Aviv Shachak[1], MSc, PhD; Emily Seto[1], MSc, PhD

[1]Department of Health Informatics, Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

[2]Department of Psychiatry, University Hospital Limerick, Limerick, Ireland

Corresponding Author:
Waqas Ullah Khan, BA, BEd, BSc, MSc, MD
Department of Health Informatics
Institute of Health Policy, Management and Evaluation
University of Toronto
155 College Street
4th Floor
Toronto, ON, M5T3M6
Canada
Phone: 1 9059037401
Email: waqas.khan@alum.utoronto.ca

Abstract

The decision to accept or reject new digital health technologies remains an ongoing challenge among health care patients, providers, technology companies, and policymakers. Over the past few decades, interest in understanding the choice to adopt technology has led to the development of numerous theories and models. In 1979, however, psychologists Kahneman and Tversky published their seminal research article that has pioneered the field of behavioral economics. They named their model the prospect theory and used it to explain decision-making behaviors under conditions of risk and uncertainty as well as to provide an understanding of why individuals may make irrational or inconsistent choices. Although the prospect theory has been used to explain decision-making in economics, law, political science, and clinically, at the individual level, its application to understanding choice in the adoption of digital health technology has not been explored. Herein, we discuss how the main components of the prospect theory's editing phase (framing effect) and evaluation phase (value function and weighting function) can provide valuable insight on why health care patients, providers, technology companies, and policymakers may decide to accept or reject digital health technologies.

(J Med Internet Res 2022;24(2):e32714) doi:10.2196/32714

KEYWORDS

decision-making; digital health technology adoption; prospect theory

The Challenge to Accept or Reject a Digital Health Technology

Digital health technology includes a broad spectrum of categories, such as mobile health, clinical information systems, wearable devices, telehealth, and personalized medicine. From mobile fitness applications to machine learning technologies used to improve disease diagnosis, clinical decision-making, and health care delivery, digital health is revolutionizing the field of medicine [1]. Nevertheless, the decision to accept or reject new digital health technologies remains an ongoing challenge among health care patients, providers, technology companies, and policymakers. Over the past few decades, interest in understanding the choice to adopt technology has led to the development of numerous theories and models.

The Technology Acceptance Model

The Technology Acceptance Model (TAM), proposed by Fred Davis in 1989, is one of the most prominent theories developed to explain the behavioral intention to use a technological innovation [2]. Applying concepts from the theory of reasoned action and theory of planned behavior, the TAM suggests that the intended use of technology is determined by its "perceived ease of use" and "perceived usefulness." Both factors are connected with one's behavioral intention, which is linked to actual behavior and system characteristics (external variables) [2,3]. Although the TAM's simplified rationale has made it

popular among researchers, academics, and within the technology industry, it is also one of its main criticisms [4]. In many studies, it is reduced to concepts that make outcomes such as intended use the primary finding, which lowers the TAM's explanatory capability and provides little insight on the actual use of the technology studied [4].

The Prospect Theory

In 1979, however, psychologists Kahneman and Tversky published their seminal research article that has pioneered the field of behavioral economics by providing a novel understanding of decision-making. They named their model the prospect theory and used it to explain decision-making behaviors under conditions of risk and uncertainty [5]. This theory also provides an understanding of why individuals may make irrational or inconsistent decisions [5]. Although the prospect theory has been used to explain decision-making in economics, law, political science, and clinically, from the physician's perspective, its application to understanding choice in the adoption of digital health technology has not been explored [6]. Notably, this can provide valuable insight on why health care patients, providers, technology companies, and policymakers may decide to accept or reject digital health technologies under conditions of risk or uncertainty as well as provide an explanation for the perceived irrational or inconsistent decisions concerning the adoption of digital health technology in medicine [7].

The Editing Phase of the Prospect Theory

The prospect theory consists of two main phases, an editing phase and an evaluation phase. In the editing phase, individuals perform a preliminary analysis of choices or prospects they are provided. This allows one to organize and edit the choices available with the aim of simplifying the decision-making process [5]. It is also during this phase that reference points, often referred to as framing effects, are made to help guide one's decision. However, framing effects can be influenced by the order, method, or wording in which they are presented [5].

Regarding the adoption of digital health technology, a framing effect can occur when a patient is offered a choice concerning their management of care. For example, individuals with a high risk of experiencing an acute cardiac event can be offered the choice to continue their current level of care (eg, routine clinic visits) or adopt a novel wearable digital health device, such as a smart watch that monitors their vital signs dynamically and advises them when to modify their lifestyle behaviors or seek medical attention.

Similar to Tversky and Kahneman's framing effect experiments involving hypothetical infectious disease management scenarios, the adoption of wearable digital health devices can be framed in various manners. For instance, the use of a wearable digital health device can be described as providing a 100% likelihood of improving one-third of users' quality of life when compared with the current standard of care. Conversely, the same device can be described as offering a 100% probability that two-thirds of users will not experience any quality-of-life improvement when compared with the current standard of care. In the first scenario, the choice is framed as an opportunity to achieve a gain of improved quality of life. Conversely, choosing the digital health device in the second example is framed as a loss as there is a greater likelihood of no benefit, but wasted time. In reality, the outcomes of both choices are the same. However, Kahneman and Tversky discovered that people make different decisions among the same choices with certainty in gains (100% likelihood of improving one-third of users' quality of life) being more desirable than certainty in loses (100% probability that two-thirds of users do not experience an improvement in their quality of life) [5].

The Evaluation Phase of the Prospect Theory

Once a choice is simplified and framed for a decision, the evaluation phase of the prospect theory begins. There are two key components in this phase, the value function and weighting function. The value function refers to the original reference point or choice, which is often the individual's current status. This baseline reference point is then compared with an operative reference point, often a future goal. An important characteristic of the value function is that the worth of each choice is dependent on how the individual perceives the change between their current and future states [5]. For example, when compared with the current standard of care (routine clinic visits), the value function for patients with cardiovascular disease using a digital health smart watch can be described as providing continuous medical feedback that fosters patient autonomy, education, and personalized preventative care while reducing health care costs.

Another important aspect of the value function is that individuals are more risk averse in the context of making gains. For example, many telehealth companies are reporting significant increases in revenue as their digital platform usage grows due to the United States' COVID-19 government policies promoting social distancing, equal reimbursement for virtual and physical visits, and the permission granted to physicians to use these tools to practice care across state lines, as well as without requiring additional state medical licenses. If these companies choose to save the additional profits or gains made during the COVID-19 pandemic instead of using them to upgrade or develop new products, they are demonstrating the value function's risk aversion approach. Conversely, Kahneman and Tversky found that individuals are more likely to take risks in the context of experiencing losses [5]. This would mean that if telehealth companies experienced substantial revenue losses during the COVID-19 pandemic, they would likely take greater financial risks to return to a favorable position [5].

The final concept of the value function is known as loss aversion, which states losses are more painful than equal gains are pleasing. In a hypothetical scenario, health care policymakers are presented with a choice to adopt a novel electronic health record (EHR) system that aims to reduce hospital costs and inefficiencies by improving resource management. However, the risk in adopting this system is a loss of $100 million of taxpayers' money if it fails. Conversely, if it succeeds, $100 million of taxpayers' money will be saved. In their research,

Kahneman and Tversky found that when individuals are presented with a similar loss or equal gains choice, they must be offered at least twice as much in gains as compared with losses to take the risk [8]. Concerning the EHR example above, this means health care policymakers would be more likely to adopt the system if the net gain or money saved is $200 million or more when compared with the risk of losing $100 million if it failed.

Regarding the evaluation phase's weighting function, Kahneman and Tversky showed that people give higher value to outcomes that are less likely to develop, while giving less importance to outcomes with a medium to high probability of occurring. Moreover, individuals are more likely to take a sure loss against a small possibility of a larger loss. This is often seen with the purchase of warranties or insurance, a sure loss, on digital health devices such as smart watches that monitor vital signs or EHR systems. In this situation, the cost of repairing the device if it breaks is higher (a small possibility of a larger loss) when compared with the minimal loss of insuring the product (a sure loss).

Conclusion

The prospect theory, as illustrated by the examples given, provides valuable insight on why health care patients, providers, technology companies, and policymakers may decide to accept or reject digital health technologies. Given the rapid growth of this industry, the medical field and its stakeholders have more choice than ever regarding the adoption of digital health tools. Nevertheless, by understanding why choices are made, we can make better informed decisions when selecting our next digital health technology prospect.

Authors' Contributions

Conception and drafting of the article was done by WUK. AS and ES critically revised and approved the final version of the article to be published.

Conflicts of Interest

None declared.

References

1. What is digital health? U.S. Food & Drug Administration. URL: https://www.fda.gov/medical-devices/digital-health-center-excellence/what-digital-health [accessed 2021-03-18]
2. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
3. Marangunić N, Granić A. Technology acceptance model: a literature review from 1986 to 2013. Univers Access Inf Soc 2014 Feb 16;14(1):81-95. [doi: 10.1007/s10209-014-0348-1]
4. Shachak A, Kuziemsky C, Petersen C. Beyond TAM and UTAUT: future directions for HIT implementation research. J Biomed Inform 2019 Dec;100:103315 [FREE Full text] [doi: 10.1016/j.jbi.2019.103315] [Medline: 31629923]
5. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. Econometrica 1979 Mar:263-292 [FREE Full text] [doi: 10.2307/1914185]
6. Verma AA, Razak F, Detsky AS. Understanding choice: why physicians should learn prospect theory. JAMA 2014 Feb 12;311(6):571-572. [doi: 10.1001/jama.2013.285245] [Medline: 24519293]
7. Croskerry P. From mindless to mindful practice--cognitive bias and clinical decision making. N Engl J Med 2013 Jun 27;368(26):2445-2448. [doi: 10.1056/NEJMp1303712] [Medline: 23802513]
8. Tversky A, Kahneman D. Advances in prospect theory: cumulative representation of uncertainty. J Risk Uncertain 1992 Oct:297-323 [FREE Full text] [doi: 10.1007/bf00122574]

Abbreviations

EHR: electronic health record
TAM: Technology Acceptance Model

Edited by A Mavragani; submitted 06.08.21; peer-reviewed by M Salimi, E Khoong, A Tannoubi, C Pearce; comments to author 22.09.21; revised version received 07.11.21; accepted 26.12.21; published 07.02.22.

Please cite as:
Khan WU, Shachak A, Seto E
Understanding Decision-Making in the Adoption of Digital Health Technology: The Role of Behavioral Economics' Prospect Theory
J Med Internet Res 2022;24(2):e32714
URL: https://www.jmir.org/2022/2/e32714
doi:10.2196/32714
PMID:35129459

Viewpoint

An Ethics Checklist for Digital Health Research in Psychiatry: Viewpoint

Francis X Shen[1,2*], JD, PhD; Benjamin C Silverman[1,3*], MD; Patrick Monette[1,3], BA; Sara Kimble[3], BS; Scott L Rauch[1,3], MD; Justin T Baker[1,3], MD, PhD

[1]Harvard Medical School, Boston, MA, United States

[2]Law School, University of Minnesota, Minneapolis, MN, United States

[3]Institute for Technology in Psychiatry, McLean Hospital, Belmont, MA, United States

[*]these authors contributed equally

Corresponding Author:
Francis X Shen, JD, PhD
Harvard Medical School
641 Huntington Ave
Boston, MA, 02115
United States
Phone: 1 617 462 3845
Email: fshen1@mgh.harvard.edu

Abstract

Background: Psychiatry has long needed a better and more scalable way to capture the dynamics of behavior and its disturbances, quantitatively across multiple data channels, at high temporal resolution in real time. By combining 24/7 data—on location, movement, email and text communications, and social media—with brain scans, genetics, genomics, neuropsychological batteries, and clinical interviews, researchers will have an unprecedented amount of objective, individual-level data. Analyzing these data with ever-evolving artificial intelligence could one day include bringing interventions to patients where they are in the real world in a convenient, efficient, effective, and timely way. Yet, the road to this innovative future is fraught with ethical dilemmas as well as ethical, legal, and social implications (ELSI).

Objective: The goal of the Ethics Checklist is to promote careful design and execution of research. It is not meant to mandate particular research designs; indeed, at this early stage and without consensus guidance, there are a range of reasonable choices researchers may make. However, the checklist is meant to make those ethical choices explicit, and to require researchers to give reasons for their decisions related to ELSI issues. The Ethics Checklist is primarily focused on procedural safeguards, such as consulting with experts outside the research group and documenting standard operating procedures for clearly actionable data (eg, expressed suicidality) within written research protocols.

Methods: We explored the ELSI of digital health research in psychiatry, with a particular focus on what we label "deep phenotyping" psychiatric research, which combines the potential for virtually boundless data collection and increasingly sophisticated techniques to analyze those data. We convened an interdisciplinary expert stakeholder workshop in May 2020, and this checklist emerges out of that dialogue.

Results: Consistent with recent ELSI analyses, we find that existing ethical guidance and legal regulations are not sufficient for deep phenotyping research in psychiatry. At present, there are regulatory gaps, inconsistencies across research teams in ethics protocols, and a lack of consensus among institutional review boards on when and how deep phenotyping research should proceed. We thus developed a new instrument, an Ethics Checklist for Digital Health Research in Psychiatry ("the Ethics Checklist"). The Ethics Checklist is composed of 20 key questions, subdivided into 6 interrelated domains: (1) informed consent; (2) equity, diversity, and access; (3) privacy and partnerships; (4) regulation and law; (5) return of results; and (6) duty to warn and duty to report.

Conclusions: Deep phenotyping research offers a vision for vastly more effective care for people with, or at risk for, psychiatric disease. The potential perils en route to realizing this vision are significant; however, and researchers must be willing to address the questions in the Ethics Checklist before embarking on each leg of the journey.

(J Med Internet Res 2022;24(2):e31146) doi:10.2196/31146

KEYWORDS

digital phenotyping; computataional psychiatry; ethics; law; privacy; informed consent

Introduction

"The deeper you go, the more you know." This headline captures the tantalizing promise of deeply probing digital health research in psychiatry [1].

Psychiatry has long needed a better and more scalable way to capture the dynamics of behavior and its disturbances, quantitatively across multiple data channels, at high temporal resolution in real time. By combining 24/7 data—on location, movement, email and text communications, and social media—with brain scans, genetics, genomics, neuropsychological batteries, and clinical interviews, researchers will have an unprecedented amount of objective, individual-level data [2,3]. Analyzing these data with ever-evolving artificial intelligence offers the possibility of intervening early with precision and could even prevent the most critical sentinel events [4]. Ideally, this could one day include bringing interventions to patients where they are in the real world, in a convenient, efficient, effective, and timely way [5,6].

Yet, the road to this innovative future is fraught with ethical dilemmas [7-10], and ethical, legal, and social implications (ELSI) on issues such as informed consent and return of results must be reexamined in light of research employing previously unavailable deep and computational characterization of humans based on: (1) real time and cross-sectional behavioral measures; (2) imaging, interviews, and other phenotypic data; (3) genotypic data; and (4) epigenetic and environmental data. The potentially boundless data collection across these streams gives rise to what might be described as "deep biobehavioral typing" or "deep geno/phenotyping." Given these multiple streams, ethical frameworks for deep biobehavioral typing must integrate the overlapping ethics of genetics, genomics, ethics of brain imaging, ethics of real time digital monitoring, and so forth to critically reexamine well-known ELSI issues. For example, when considering return of results, although it is true that the deeper you go, the more you know, it is unclear when a researcher knows enough to justify sharing data with a clinician, alerting appropriate individuals about potential self-harm, and returning individual research results [11].

Supported by a National Institutes of Health (NIH) Bioethics Administrative Supplement award (NIH 1U01MH116925-01), we have been exploring the ELSI of digital health research in psychiatry, with a particular focus on what we label "deep phenotyping" psychiatric research, which combines the potential for virtually boundless data collection and increasingly sophisticated techniques to analyze that data. We convened an interdisciplinary expert stakeholder workshop in May 2020, and this checklist emerges out of that dialogue. As we use it in this article, the phrase "deep phenotyping" in psychiatric research is meant to describe research that—even if it does not encompass a large number of research subjects—goes deep into the lives of those subjects by collecting many digital and biological data streams (eg, digital data such as text messages, phone screen shots, and GPS location; health data such as heart rate and blood pressure; and clinical evaluations and biological data such as genetics and brain scans).

Consistent with recent ELSI analyses [8,9], the bottom line of our bioethics analysis is that existing ethical guidance and legal regulation are not sufficient for deep phenotyping research in psychiatry. At present, there are regulatory gaps and inconsistencies across research teams in ethics protocols. There is also a lack of consensus among institutional review boards (IRBs) on when and how deep phenotyping research should proceed [12,13]. Efforts are underway to fill these gaps, notably those led by the Connected and Open Research Ethics initiative at the University of California San Diego [9].

Until the field develops more robust consensus guidelines, however, the onus clearly falls on individual research teams to take the lead in shaping the applied ethics of digital health research in psychiatry.

To guide these ethics considerations, we developed a new instrument, an Ethics Checklist for Digital Health Research in Psychiatry ("the Ethics Checklist"). The Ethics Checklist is composed of 20 key questions, subdivided into six interrelated domains: (1) informed consent; (2) equity, diversity, and access; (3) privacy and partnerships; (4) regulation and law; (5) return of results; and (6) duty to warn and duty to report. The questions included in the checklist are presented in Table 1, and Multimedia Appendix 1 provides the Ethics Checklist as a user-friendly instrument for research teams.

The goal of the Ethics Checklist is to promote the careful design and execution of research. It is not meant to mandate particular research designs; indeed, at this early stage of digital phenotyping research and without consensus guidance, there are a range of reasonable choices researchers may make. But the checklist is meant to make those ethical choices explicit, and to require researchers to give reasons for their decisions related to ELSI issues. The Ethics Checklist is primarily focused on procedural safeguards, such as consulting with experts outside the research group and documenting standard operating procedures for clearly actionable data (eg, expressed suicidality) within written research protocols.

Table 1. Ethics checklist for digital health research.

Category	Category description	Checklist items
Informed consent	How can we meaningfully communicate and be transparent about research methods that involve deep, complex, often passive and continuous data collection, machine learning analysis, and interpretation?	1. Have we appropriately adapted our informed consent procedures to our specific study population, including possible use of surrogate consent?
		2. Will we provide background education on relevant technologies, such as explaining what social media companies may already be doing with the participant's data?
		3. Have we determined what a reasonable person would want to know, and explained in our institutional review board proposal the evidence on which we reached that determination?
Equity, diversity, and access	How will we address concerns that our research might replicate existing, or generate new, biased results or contribute to health inequities in access based on race, ethnicity, gender, sexual orientation, age, or another legally protected class?	4. Starting at the early conceptualization and research design stages, have we sought input from a diverse community of stakeholders to identify and address potential equity concerns and opportunities to advance justice with our proposed research?
		5. Has our research plan addressed potential inequities in access, for instance varying levels of access to mobile technology and to health care services?
		6. Has every member of the research team completed our institution's recommended trainings around diversity, inclusion, equity, and access?
Privacy and partnerships	How can we design our research to balance an interest in robust data collection, with a potentially competing interest in protecting participant privacy?	7. Have we consulted with information security experts about exactly where the data will flow, from start to finish?
		8. Do we have a written policy on data deidentification and participant privacy that is consistent with best practices in psychiatry and neuroscience?
		9. Have we determined which, if any, third-party vendors will be required to be HIPAA[a] compliant and sign a Business Associate Agreement?
Regulation and law	Which state, federal, and international law and regulatory guidance must be adhered to in our research?	10. Have we examined the terms of service, end user license agreements, privacy statements, and HIPAA notices for each of the vendors and software applications involved in our research?
		11. Have we determined how laws in applicable jurisdictions will treat the data we collect, for instance considering the data to be "sensitive," "special category," or "personal health information"?
		12. Have we ensured compliance with state, federal and international laws governing our research, HIPAA privacy requirements, state data privacy laws, and applicable international privacy laws?
Return of results	By which criteria will we determine if our data analytic models are sufficiently valid and reliable for us to share the individual research results and data with the research participant and the participant's clinicians?	13. Have we considered whether our study will generate any "actionable" results, based on established guidelines and how we have defined actionability?
		14. Have we established with what frequency results will be returned? (eg, should participants have daily, weekly, and monthly access to some subset of their data?)
		15. Have we clarified the protocols and mechanisms for returning different types of information, (eg, raw data, interpreted data, etc)?
		16. Do we have a protocol in place for contacting a participant's clinicians and nonclinical caregivers?

Category	Category description	Checklist items
Duty to warn and duty to report	When might our research trigger a legal or ethical duty to report the potential for participant self-harm or harm to others, and what are our protocols for determining whether in individual instances we have such a duty?	17. Has everyone in our research lab received sufficient training to know when to flag data or results as requiring follow-up review by a supervisor? 18. Will our analytic methods allow us to identify the precursors to dangerous or illegal behavior, to oneself or to others, and if so, at which point will we intervene to protect the research participant or a third party? 19. Have we updated our lab's suicidality standard operating procedure to be consistent with the novel data acquisition and analysis techniques we are using in our study? 20. Do we have a protocol for responding to legally mandated reporting if our data uncover child pornography, restraining order violations, and so on?

[a]HIPAA: Health Insurance Portability and Accountability Act.

The Ethics Checklist in Action

Each of the 20 ethics checklist questions are phrased so that they can be answered with a "Yes," "No," or "Pending" response. In our view, deep phenotyping research in psychiatry should not proceed until a research team answers "Yes" or "Pending" to each checklist question. To arrive at "Yes" or "Pending" for each question will require research labs to carefully consider a complex interplay of ethical and legal considerations.

It is beyond the scope of this short paper to address all of these complexities, but we offer here several illustrative examples, from each of the 6 key domains, of how the checklist might be applied in practice.

Informed Consent

The revised Common Rule requires researchers to present participants with information that is "most likely to assist … in understanding the reasons why one might or might not want to participate in the research," and that is what "a reasonable person would want to have …" (Title 45 of the Code of Federal Regulations, part 46, effective July 2018) [14]. The Ethics Checklist thus requires researchers to address the question: "Have we determined what a reasonable person would want to know, and explained in our IRB proposal the evidence on which we reached that determination?" There are presently no empirical data or standard protocols for determining what information a reasonable person would want to have before agreeing to participate in deep phenotyping research. Privacy scholars and ethicists are increasingly concerned that the rights of "notice, access, and consent regarding the collection, use, and disclosure of personal data" are no longer adequate because "many privacy harms are the result of an aggregation of pieces of data over a period of time by different entities" [15]. Thus, researchers must consider a broad range of possible information to communicate.

We offer here several of these many possibilities. One decision is whether to provide research participants a list of clear, concrete examples of the inferences that can likely be made from participants' data. For example, the informed consent material could explicitly say, "You should know that, although we will not reveal this information outside the research team, we may be able to identify when you are going to the bathroom or having sex." Another decision, especially for researchers who are collecting data on participants' GPS data and social media content, is whether to provide basic background education to make participants more informed on the data collection and data sharing practices already being utilized by the mobile technology and apps they already use regularly. Third, ethics research has identified a need to make informed consent processes more meaningful and valid by improving communication [16,17]. Research teams may consider innovative strategies such as video-based multimedia as part of the consent procedure [18-20]. They may also consider staged informed consent [21,22], dynamic consent to facilitate 2-way communication between researchers and participants [23], or a systemic oversight approach for big data research that provides flexibility for addressing uncertainty in how data will be used [24].

In addition, researchers in psychiatry must address a further question that has long been challenging for the field, "how to ensure meaningful and valid informed consent with participants who have a mental illness?" [25,26]. The question of decisional capacity in psychiatric research has been well researched, with multiple instruments now available [27], but the field will need to revisit the effectiveness of these instruments in the new context of deep biobehavioral research.

Equity, Diversity, Inclusion, and Access

It is well established that biomedical research generally [28] and psychiatric research specifically struggle to enroll racially and ethnically diverse participants [29]. In the United States, for example, there is a reluctance of African Americans to participate in research given a long history of racism and exploitation [30,31].

The Ethics Checklist proposes that researchers answer the following question: Starting at the early conceptualization and research design stages, have we sought input from a diverse community of stakeholders to identify and address potential equity concerns and opportunities to advance justice with our proposed research? The question emphasizes that equity concerns extend beyond simply developing proportional samples, and that from the start, "[r]esearch relationships must

become balanced, reciprocal, and community informed, without centering researcher and institutional priorities" [32]. Community advisory boards, which can be comprised of community and family stakeholders, can be a useful vehicle for facilitating such engagement [33]. Operationalizing "diverse community of stakeholders" will depend on the nature and scope of the research, institutional context, and affected communities. The stakeholder group will necessarily be comprised differently; for instance, whether the focus is on a single disease versus basic research, or the research involves multiple international sites versus a single community partner.

In defining the stakeholders, the checklist encourages researchers to go beyond their own research team to seek guidance and build trust, even at the conceptual and research design stage. We agree with Wilkins [34], who suggests that enhancing trust "must build on the principles of community engagement including balancing power dynamics, equitable distribution of resources, effective bidirectional communication, shared decision-making, and valuing of different resources and assets (such as the lived experience and knowledge of group norms and perspectives)." Researchers might, for instance, consult with their institution's leadership on diversity and inclusion to see if the institution already has mechanisms in place for community engagement. Additional options include reviewing the best practices in community-based participatory research and community-engaged research [35] and consulting expertise in other disciplines, including law [36] and the humanities [37].

Privacy and Partnerships

In an analysis of smartphone digital phenotyping, Onnela and Rauch [38] write that "[p]atient and participant privacy is always of utmost importance in clinical and research settings." The literature on ethics of deep phenotyping has similarly identified privacy as foundational [39,40]. Deep phenotyping research requires that data flow across multiple platforms and vendors; thus, to safeguard data privacy, researchers must be aware of where the data go, what happens to the data at each stop, and where security vulnerabilities may exist [10]. The Ethics Checklist thus includes the question, "Have we consulted with information security experts about exactly where the data will flow, from start to finish?" Such consultation could potentially lead to modifications in data collection and data analysis techniques to improve privacy safeguards. For instance, security experts may be aware of the vulnerabilities of particular apps or new technical advances recently developed.

Regulation and Law

The regulation of mobile health apps is currently undergoing transformation [41,42], as is the regulation of artificial intelligence and machine learning data analysis [43]. At the same time, state privacy laws are emerging [44], as are international law innovations such as the European Union's General Data Protection Regulation [45]. These legal developments have implications for the deep phenotyping research we describe in this article [7].

For instance, the data collection may require interfacing with multiple third-party vendors, and it is the responsibility of the research team to examine the terms of service, end user license agreements, privacy statements, and Health Insurance Portability and Accountability Act (HIPAA) notices for each of these vendors and associated software applications (Checklist question 10). This may not be an easy task, as research on mobile health apps suggests that many vendors do not have a privacy policy publicly available [46]. It will also be challenging to determine how applicable laws will treat the data being collected (Checklist question 11). Different laws define categories differently. For example, what is considered "sensitive" data under the California Consumer Privacy Act and California Privacy Rights Act might be different from what is considered "special category" data under the European Union's General Data Protection Regulation or considered "personal health information" under HIPAA [47]. In setting up the research design, the research team may need to enlist institutional or external expertise to help understand and comply with these statutes.

In addition, when data collection follows the individual across state or international boundaries, and when data flow across those boundaries, the research will be exposed to multiple legal jurisdictions, including emerging state laws governing privacy and research [48]. For example, if data is collected continuously while a research participant living in Boston visits a relative in Chicago, then goes to a meeting in Baltimore, both the Illinois Biometric Privacy Act and the Maryland Confidentiality of Medical Records Act may apply. This may place different requirements on the research team. Similarly, if data gathered from a research participant in Detroit are transferred to a vendor operating in nearby Windsor, Canada's data privacy laws may now be relevant in a way they would not be for traditional research with subjects and data firmly rooted in the United States. Given this potential for movement of research subjects, researchers should answer the question, "Have we ensured compliance with state, federal and international laws governing our research, HIPAA privacy requirements, state data privacy laws, and applicable international privacy laws?" Reviewing HIPAA compliance is foundational, and the details of such review are beyond the scope of our commentary; however, we emphasize here that, given the geographic mobility of the subjects, HIPAA is not the only applicable privacy regime. Thus, addressing legal and regulatory concerns may likely require consultation with legal experts in the researcher's institution. This review of privacy law can be integrated with the Ethics Checklist question 10 concerning vendors' policies and question 11 concerning the designation of sensitive information.

Return of Results

The return of individual research results has garnered significant attention in the ethics literature [49-51]. In the context of deep phenotyping research, the return of results is challenging because there are potentially so much data to return and because some of that data fluctuate over time and could be returned hourly, daily, weekly, monthly, and so on. In a virtual workshop we hosted in May 2020, a group of 25 stakeholders from science, medicine, law, and ethics gathered to explore the issues of return of results in deep phenotyping research. The discussion in that workshop made clear that, at present, there is considerable

uncertainty over what constitutes "actionable" information in this space as well as divergent practices among research teams in which information participants can access. The Ethics Checklist includes 4 different questions on return of results, including, "Have we considered whether our study will generate any 'actionable' results, based on established guidelines, and how have we defined actionability?" The concept of actionability is debated across multiple fields [50,52], and in the context of deep phenotyping it is not clear where the bounds of actionability are.

For instance, if a research team is measuring step count data and a participant's step count drops below average in a given week, alerting that participant of the data is actionable in the sense that the participant—informed by these data—may choose to walk substantially more steps next week. But what about more complex results, such as a machine learning algorithm that predicts that the participant has a 72% higher likelihood of experiencing a manic episode in the following year? When has the scientific knowledge base accumulated sufficiently to make such a prediction "actionable"? An even more fundamental question is implicated: for any measurement or prediction, what is the confidence in the measurement, sensitivity, or specificity of the interpretation or prediction, and how should that be shared? The effects of researcher mobile health interventions on participants within research studies are only now being studied [53], and the field is still formulating practices for clinical interventions based on phenotyping data [54].

The potential ethical responsibility and legal duty for reanalysis of data is also of concern. For instance, in genomics research, many genetic variations are classified as "variant of uncertain significance (VUS)," but as knowledge increases, those variations may be reclassified [55]. An ethical and legal question is whether researchers should (or must) revisit previously collected data to determine whether reclassification is warranted, and if so, when and how should they contact those participants from the earlier study [56]. This issue will likely emerge in deep geno/phenotyping research, and it should be proactively anticipated with a policy put into place.

Duty to Warn and Duty to Report

The duty to warn and the duty to report are well known to psychiatric researchers, but the advanced data collection and data analysis methods of deep phenotyping introduce unique concerns [57]. At present, the field has only begun to develop protocols and thresholds for when the data should trigger a legal or ethical duty to report the potential for participants' self-harm or harm to others, and further work is needed to address the possibility of false positives, false negatives, and reliable signal detection. For instance, among the 4 questions included in the Ethics Checklist under this heading, we require that researchers address the question, "Have we updated our lab's suicidality standard operating procedure (SOP) to be consistent with the novel data acquisition and analysis techniques we are using in our study?" Traditionally, suicidality SOPs have relied almost exclusively on the clinical judgment of the psychiatrist or psychologist reviewing individual records and conducting interviews with the participant. The goal of deep phenotyping research, however, is to reduce reliance on this single stream of data, and instead to incorporate many additional real-world data points. The suicidality SOP may need to be modified in recognition of this new paradigm of psychiatric assessment. For instance, if GPS data show that the participant has spent 3 hours at a local bar, then is located on a bridge at 2 AM in the morning, and has sent 20 text messages in the past 5 minutes, is there a way for the research team to have such behavior flagged in real time and should there be a real time intervention in the protocol? The Ethics Checklist requires that researchers consider such situations.

The Urgent Need for Consensus Guidance

Deep phenotyping research offers a vision for vastly more effective care for people with or at risk of psychiatric disease. The potential perils en route to realizing this vision are significant; however, researchers must be willing to address the questions in the Ethics Checklist before embarking on each leg of the journey.

The illustrative examples discussed above make clear that deep phenotyping researchers have few guideposts and little empirical data with which to address many pressing ethical and legal questions critical for their research. This lack of clarity regarding best practices is understandable for a field that has emerged rapidly, mainly in the past 5 years. But as the field continues to expand, there is a need to fill this gap by developing consensus guidance, informed by quantitative and qualitative bioethics research, as well as community and patient advocate input. This paper has raised more questions than answers, and it did not reach many other avenues of inquiry including considerations for international research and research with children.

To make progress toward consensus guidance, we identify 2 immediate action items. First, ethics should be integrated into the practice of deep phenotyping research (as is already being carried out at centers such as the McLean Institute for Technology in Psychiatry and the Connected and Open Research Ethics initiative at the University of California San Diego Research Center for Optimal Digital Ethics in Health).

Second, professional organizations such as the American Psychiatric Association and the Digital Medicine Society, along with institutions such as the NIH and National Academies, are well positioned to convene an interdisciplinary team to conduct in-depth analysis and produce foundational reports to guide the field.

The deeper you go in deep phenotyping research, the deeper the ethical and legal challenges. But with timely, concerted action, the research community can promote ethically sound and legally compliant digital health research in psychiatry.

Acknowledgments

We thank our student research assistant Lois Yoo. We also thank the participants in a virtual workshop hosted by the McLean Institute for Technology in Psychiatry on May 8, 2020, to explore the ethical, legal, and social implications of return of results in deep phenotyping research. Research reported in this publication was supported by a Bioethics Supplement from the National Institute of Mental Health (NIMH) of the National Institutes of Health (NIH) under Award Number 1U01MH116925-01. The content is solely the responsibility of the authors and does not necessarily represent the official views of NIMH or NIH.

Authors' Contributions

FXS, BCS, and JTB conceived and designed the work. FS and BS substantially conducted the ethics research and drafted the manuscript. SLR and JTB substantially edited and critically revised the manuscript. PM and SK significantly contributed to the ethics research and revised the manuscript. All the authors gave final approval of the completed manuscript version and are accountable for all aspects of the work.

Conflicts of Interest

SLR is employed by McLean Hospital/Mass General Brigham, serves as the secretary of Society of Biological Psychiatry, and on the Boards of McLean Hospital, National Network of Depression Centers, National Association of Behavioral Healthcare, and Community Psychiatry/Mindpath Health. He receives royalties from Oxford University Press and APPI. None of these are known to represent conflicts of interest with this work. JTB has received consulting fees from Verily Life Sciences, as well as consulting fees and equity from Mindstrong Health, Inc, unrelated to the present work.

Multimedia Appendix 1
Ethics checklist for digital health research in psychiatry and worksheet.
[DOCX File , 30 KB - jmir_v24i2e31146_app1.docx]

References

1. Delude CM. Deep phenotyping: The details of disease. Nature 2015 Nov 05;527(7576):S14-S15. [doi: 10.1038/527S14a] [Medline: 26536218]
2. Baker JT, Germine LT, Ressler KJ, Rauch SL, Carlezon WA. Digital devices and continuous telemetry: opportunities for aligning psychiatry and neuroscience. Neuropsychopharmacology 2018 Dec;43(13):2499-2503 [FREE Full text] [doi: 10.1038/s41386-018-0172-z] [Medline: 30120409]
3. Torous J, Onnela JP, Keshavan M. New dimensions and new tools to realize the potential of RDoC: digital phenotyping via smartphones and connected devices. Transl Psychiatry 2017 Mar 07;7(3):e1053 [FREE Full text] [doi: 10.1038/tp.2017.25] [Medline: 28267146]
4. Sanislow CA, Ferrante M, Pacheco J, Rudorfer MV, Morris SE. Advancing Translational Research Using NIMH Research Domain Criteria and Computational Methods. Neuron 2019 Mar 06;101(5):779-782 [FREE Full text] [doi: 10.1016/j.neuron.2019.02.024] [Medline: 30844398]
5. Insel TR. Digital Phenotyping: Technology for a New Science of Behavior. JAMA 2017 Oct 03;318(13):1215-1216. [doi: 10.1001/jama.2017.11295] [Medline: 28973224]
6. Baker JT. The Digital Future of Psychiatry. Psychiatric Annals 2019 May;49(5):193-194. [doi: 10.3928/00485713-20190416-02]
7. Martinez-Martin N, Insel TR, Dagum P, Greely HT, Cho MK. Data mining for health: staking out the ethical territory of digital phenotyping. NPJ Digit Med 2018;1:1-5 [FREE Full text] [doi: 10.1038/s41746-018-0075-8] [Medline: 31211249]
8. Nebeker C, Bartlett Ellis RJ, Torous J. Development of a decision-making checklist tool to support technology selection in digital health research. Transl Behav Med 2020 Oct 08;10(4):1004-1015 [FREE Full text] [doi: 10.1093/tbm/ibz074] [Medline: 31120511]
9. Torous J, Nebeker C. Navigating Ethics in the Digital Age: Introducing Connected and Open Research Ethics (CORE), a Tool for Researchers and Institutional Review Boards. J Med Internet Res 2017 Feb 08;19(2):e38 [FREE Full text] [doi: 10.2196/jmir.6793] [Medline: 28179216]
10. Kaplan B. Revisiting Health Information Technology Ethical, Legal, and Social Issues and Evaluation: Telehealth/Telemedicine and COVID-19. Int J Med Inform 2020 Nov;143:104239 [FREE Full text] [doi: 10.1016/j.ijmedinf.2020.104239] [Medline: 33152653]
11. Ienca M, Ferretti A, Hurst S, Puhan M, Lovis C, Vayena E. Considerations for ethics review of big data health research: A scoping review. PLoS One 2018;13(10):e0204937 [FREE Full text] [doi: 10.1371/journal.pone.0204937] [Medline: 30308031]
12. Torous J, Wisniewski H, Bird B, Carpenter E, David G, Elejalde E, et al. Creating a Digital Health Smartphone App and Digital Phenotyping Platform for Mental Health and Diverse Healthcare Needs: an Interdisciplinary and Collaborative Approach. J. technol. behav. sci 2019 Apr 27;4(2):73-85. [doi: 10.1007/s41347-019-00095-w]

13. Ball TM, Kalinowski A, Williams LM. Ethical implementation of precision psychiatry. Personalized Medicine in Psychiatry 2020 Mar 20;19-20:100046. [doi: 10.1016/j.pmip.2019.05.003]
14. 2018 Requirements (2018 Common Rule): Code of Federal Regulations Title 45 Public Welfare Department Of Health And Human Services Part 46: Protection Of Human Subjects. Office for Human Research Protections. URL: https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/revised-common-rule-regulatory-text/index.html [accessed 2022-01-25]
15. Solove DJ. Privacy self-management and the consent dilemma. Harvard Law Review 2012 Nov 04;126:1880-1903.
16. Kadam RA. Informed consent process: A step further towards making it meaningful!. Perspect Clin Res 2017;8(3):107-112 [FREE Full text] [doi: 10.4103/picr.PICR_147_16] [Medline: 28828304]
17. Wolf SM, Clayton EW, Lawrenz F. The Past, Present, and Future of Informed Consent in Research and Translational Medicine. J. Law. Med. Ethics 2021 Jan 01;46(1):7-11. [doi: 10.1177/1073110518766003]
18. Synnot A, Ryan R, Prictor M, Fetherstonhaugh D, Parker B. Audio-visual presentation of information for informed consent for participation in clinical trials. Cochrane Database Syst Rev 2014 May 09(5):CD003717 [FREE Full text] [doi: 10.1002/14651858.CD003717.pub3] [Medline: 24809816]
19. Anderson EE, Newman SB, Matthews AK. Improving informed consent: Stakeholder views. AJOB Empir Bioeth 2017;8(3):178-188 [FREE Full text] [doi: 10.1080/23294515.2017.1362488] [Medline: 28949896]
20. Kraft SA, Constantine M, Magnus D, Porter KM, Lee SS, Green M, et al. A randomized study of multimedia informational aids for research on medical practices: Implications for informed consent. Clin Trials 2017 Feb;14(1):94-102 [FREE Full text] [doi: 10.1177/1740774516669352] [Medline: 27625314]
21. Young-Afat DA, Verkooijen HAM, van Gils CH, van der Velden JM, Burbach JP, Elias SG, et al. Brief Report: Staged-informed Consent in the Cohort Multiple Randomized Controlled Trial Design. Epidemiology 2016 May;27(3):389-392. [doi: 10.1097/EDE.0000000000000435] [Medline: 27035689]
22. Vickers AJ, Young-Afat DA, Ehdaie B, Kim SY. Just-in-time consent: The ethical case for an alternative to traditional informed consent in randomized trials comparing an experimental intervention with usual care. Clin Trials 2018 Feb;15(1):3-8 [FREE Full text] [doi: 10.1177/1740774517746610] [Medline: 29224379]
23. Budin-Ljøsne I, Teare HJA, Kaye J, Beck S, Bentzen HB, Caenazzo L, et al. Dynamic Consent: a potential solution to some of the challenges of modern biomedical research. BMC Med Ethics 2017 Jan 25;18(1):4 [FREE Full text] [doi: 10.1186/s12910-016-0162-9] [Medline: 28122615]
24. Vayena E, Blasimme A. Health Research with Big Data: Time for Systemic Oversight. J Law Med Ethics 2018 Mar;46(1):119-129 [FREE Full text] [doi: 10.1177/1073110518766026] [Medline: 30034208]
25. Carpenter WT, Gold JM, Lahti AC, Queern CA, Conley RR, Bartko JJ, et al. Decisional capacity for informed consent in schizophrenia research. Arch Gen Psychiatry 2000 Jun;57(6):533-538. [doi: 10.1001/archpsyc.57.6.533] [Medline: 10839330]
26. Roberts LW. Informed consent and the capacity for voluntarism. Am J Psychiatry 2002 May;159(5):705-712. [doi: 10.1176/appi.ajp.159.5.705] [Medline: 11986120]
27. Dunn LB, Nowrangi MA, Palmer BW, Jeste DV, Saks ER. Assessing decisional capacity for clinical research or treatment: a review of instruments. Am J Psychiatry 2006 Aug;163(8):1323-1334. [doi: 10.1176/ajp.2006.163.8.1323] [Medline: 16877642]
28. Oh SS, Galanter J, Thakur N, Pino-Yanes M, Barcelo NE, White MJ, et al. Diversity in Clinical and Biomedical Research: A Promise Yet to Be Fulfilled. PLoS Med 2015 Dec;12(12):e1001918 [FREE Full text] [doi: 10.1371/journal.pmed.1001918] [Medline: 26671224]
29. Brown G, Marshall M, Bower P, Woodham A, Waheed W. Barriers to recruiting ethnic minorities to mental health research: a systematic review. Int J Methods Psychiatr Res 2014 Mar;23(1):36-48 [FREE Full text] [doi: 10.1002/mpr.1434] [Medline: 24474683]
30. Gordon-Achebe K, Hairston DR, Miller S, Legha R, Starks S. Origins of Racism in American Medicine and Psychiatry. In: Medlock M, Shtasel D, Trinh NH, Williams D, editors. Racism and Psychiatry. Cham, Switzerland: Humana Press; Nov 21, 2018:3-19.
31. Huang H, Coker AD. Examining Issues Affecting African American Participation in Research Studies. Journal of Black Studies 2008 May 01;40(4):619-636. [doi: 10.1177/0021934708317749]
32. Gilmore-Bykovskyi A, Jackson JD, Wilkins CH. The Urgency of Justice in Research: Beyond COVID-19. Trends Mol Med 2021 Feb;27(2):97-100 [FREE Full text] [doi: 10.1016/j.molmed.2020.11.004] [Medline: 33277159]
33. Quinn SC. Ethics in public health research: protecting human subjects: the role of community advisory boards. Am J Public Health 2004 Jun;94(6):918-922. [doi: 10.2105/ajph.94.6.918] [Medline: 15249289]
34. Wilkins WH. Effective Engagement Requires Trust and Being Trustworthy. Med Care 2018 Oct;56 Suppl 10 Suppl 1:S6-S8 [FREE Full text] [doi: 10.1097/MLR.0000000000000953] [Medline: 30015725]
35. Wallerstein N, Calhoun K, Eder M, Kaplow J, Wilkins CH. Engaging the Community: Community-Based Participatory Research and Team Science. In: Hall K, Vogel A, Croyle R, editors. Strategies for Team Science Success. Cham, Switzerland: Springer; 2019:123-134.
36. Jones OD, Schall JD, Shen FX. Law and Neuroscience, Second Edition. Philadelphia, PA, US: Wolters Kluwer; Sep 15, 2020.

37. Boniolo G, Campaner R, Coccheri S. Why include the humanities in medical studies? Intern Emerg Med 2019 Oct;14(7):1013-1017. [doi: 10.1007/s11739-019-02131-2] [Medline: 31227997]
38. Onnela J, Rauch SL. Harnessing Smartphone-Based Digital Phenotyping to Enhance Behavioral and Mental Health. Neuropsychopharmacology 2016 Jun;41(7):1691-1696 [FREE Full text] [doi: 10.1038/npp.2016.7] [Medline: 26818126]
39. Price WN, Cohen IG. Privacy in the age of medical big data. Nat Med 2019 Jan;25(1):37-43 [FREE Full text] [doi: 10.1038/s41591-018-0272-7] [Medline: 30617331]
40. Rosenfeld L, Torous J, Vahia IV. Data Security and Privacy in Apps for Dementia: An Analysis of Existing Privacy Policies. Am J Geriatr Psychiatry 2017 Aug;25(8):873-877. [doi: 10.1016/j.jagp.2017.04.009] [Medline: 28645535]
41. Cortez NG, Cohen IG, Kesselheim AS. FDA Regulation of Mobile Health Technologies. N Engl J Med 2014 Jul 24;371(4):372-379. [doi: 10.1056/nejmhle1403384]
42. Armontrout JA, Torous J, Cohen M, McNiel DE, Binder R. Current Regulation of Mobile Mental Health Applications. J Am Acad Psychiatry Law 2018 Jun;46(2):204-211. [doi: 10.29158/JAAPL.003748-18] [Medline: 30026399]
43. Marks M. Emergent Medical Data: Health Information Inferred by Artificial Intelligence. UC Irvine Law Review 2021;11(4):995-1066 [FREE Full text]
44. Noordyke M. US State Comprehensive Privacy Law Comparison. International Association of Privacy Professionals. 2020. URL: https://iapp.org/resources/article/state-comparison-table/ [accessed 2021-01-04]
45. Mostert M, Koomen B, van Delden J, Bredenoord A. Privacy in Big Data psychiatric and behavioural research: A multiple-case study. Int J Law Psychiatry 2018;60:40-44. [doi: 10.1016/j.ijlp.2018.07.002] [Medline: 30217329]
46. Sunyaev A, Dehling T, Taylor PL, Mandl KD. Availability and quality of mobile health app privacy policies. J Am Med Inform Assoc 2015 Apr;22(e1):e28-e33. [doi: 10.1136/amiajnl-2013-002605] [Medline: 25147247]
47. Mulgund P, Mulgund BP, Sharman R, Singh R. The implications of the California Consumer Privacy Act (CCPA) on healthcare organizations: Lessons learned from early compliance experiences. Health Policy and Technology 2021 Sep;10(3):100543. [doi: 10.1016/j.hlpt.2021.100543]
48. Tovino SA. Mobile Research Applications and State Research Laws. J Law Med Ethics 2020 Mar;48(1_suppl):82-86. [doi: 10.1177/1073110520917032] [Medline: 32342744]
49. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Health Sciences Policy, Committee on the Return of Individual-Specific Research Results Generated in Research Laboratories. In: Downey AS, Busta ER, Mancher M, Botkin JR, editors. Returning Individual Research Results to Participants: Guidance for a New Research Paradigm. Washington DC, US: National Academies Press; Jul 10, 2018.
50. Wolf SM. Return of individual research results and incidental findings: facing the challenges of translational science. Annu Rev Genomics Hum Genet 2013;14:557-577 [FREE Full text] [doi: 10.1146/annurev-genom-091212-153506] [Medline: 23875796]
51. Evans BJ, Wolf SM. A Faustian Bargain That Undermines Research Participants' Privacy Rights and Return of Results. Florida Law Review 2019;71(5):1281-1345.
52. Jarvik GP, Amendola LM, Berg JS, Brothers K, Clayton EW, Chung W, eMERGE Act-ROR CommitteeCERC Committee, CSER Act-ROR Working Group, et al. Return of genomic results to research participants: the floor, the ceiling, and the choices in between. Am J Hum Genet 2014 Jun 05;94(6):818-826 [FREE Full text] [doi: 10.1016/j.ajhg.2014.04.009] [Medline: 24814192]
53. NeCamp T, Sen S, Frank E, Walton MA, Ionides EL, Fang Y, et al. Assessing Real-Time Moderation for Developing Adaptive Mobile Health Interventions for Medical Interns: Micro-Randomized Trial. J Med Internet Res 2020 Mar 31;22(3):e15033 [FREE Full text] [doi: 10.2196/15033] [Medline: 32229469]
54. Vaidyam A, Halamka J, Torous J. Actionable digital phenotyping: a framework for the delivery of just-in-time and longitudinal interventions in clinical healthcare. Mhealth 2019;5:25 [FREE Full text] [doi: 10.21037/mhealth.2019.07.04] [Medline: 31559270]
55. Roberts JL, Foulkes AL. Genetic Duties. William & Mary Law Review 2020 Nov 21;62(1):143.
56. Lu CY, Hendricks-Sturrup RM, Mazor KM, McGuire AL, Green RC, Rehm HL. The case for implementing sustainable routine, population-level genomic reanalysis. Genet Med 2020 Apr;22(4):815-816. [doi: 10.1038/s41436-019-0719-3] [Medline: 31831882]
57. Kaplan B. Seeing Through Health Information Technology: The Need for Transparency in Software, Algorithms, Data Privacy, and Regulation. SSRN Journal 2020;7(1):62. [doi: 10.2139/ssrn.3672395]

Abbreviations

ELSI: ethical, legal, and social implications
HIPAA: Health Insurance Portability and Accountability Act
IRB: institutional review board
NIH: National Institutes of Health
SOP: standard operating procedure

Edited by R Kukafka; submitted 10.06.21; peer-reviewed by N Martinez-Martin, R Hendricks-Sturrup; comments to author 28.06.21; revised version received 23.08.21; accepted 29.10.21; published 09.02.22.

Please cite as:
Shen FX, Silverman BC, Monette P, Kimble S, Rauch SL, Baker JT
An Ethics Checklist for Digital Health Research in Psychiatry: Viewpoint
J Med Internet Res 2022;24(2):e31146
URL: https://www.jmir.org/2022/2/e31146
doi:10.2196/31146
PMID:35138261

Viewpoint

Innovation Centers in Health Care Delivery Systems: Structures for Success

Onil Bhattacharyya[1,2,3], MD, PhD; Justin Shapiro[4,5], MSc; Eric C Schneider[6,7], MD, MSc

[1]Women's College Research Institute, Women's College Hospital, Toronto, ON, Canada
[2]Department of Family and Community Medicine, University of Toronto, Toronto, ON, Canada
[3]Department of Family Medicine, Women's College Hospital, Toronto, ON, Canada
[4]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada
[5]Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada
[6]Department of Health Policy and Management, T H Chan School of Public Health, Harvard University, Cambridge, MA, United States
[7]The Commonwealth Fund, New York, NY, United States

Corresponding Author:
Onil Bhattacharyya, MD, PhD
Women's College Research Institute
Women's College Hospital
76 Grenville Street
Toronto, ON, M5S 1B2
Canada
Phone: 1 416 323 6400
Email: onil.bhattacharyya@wchospital.ca

Abstract

The need to support innovation in health care delivery was prompted by payment reforms and access to digital tools and has been accelerated by the shift to virtual care as part of the COVID-19 pandemic response. Prior to the pandemic, a growing number of health systems set up innovation centers to focus on creating new services and exploring new business models relevant to value-based care. This is distinct from process improvement or implementation science, and often needs a different set of incentives to succeed within a large organization. We used a national survey to identify a diverse sample of innovation centers, and interviewed leaders to describe their aims, organizational structures, and activities. They all aim to improve patient outcomes and experience while reducing costs, but their strategic focus may differ. The centers also vary in their reporting structure, how they build internal capacity, and how they measure success. We highlight the range of strategies through examples of projects that improve quality, reduce costs, and generate new revenue. While the optimal forms and impact of innovation centers are still emerging, the fiscal pressures and the rapid uptake of digital technologies present opportunities for the redesign of health services in the postpandemic era. The experiences of these centers illustrate a set of approaches to increase any organization's capacity for innovation.

(J Med Internet Res 2022;24(2):e33961) doi:10.2196/33961

KEYWORDS

innovation; digital health; value-based care; quality improvement; delivery science; value; structure; success; health care delivery; reform; survey; outcome; experience; strategy

Introduction

While improvement and innovation are ubiquitous terms, the COVID-19 pandemic response has made innovation more of a core requirement than a marketing tool in health care. Both aim to solve problems, improve outcomes, and reduce costs, but there are key distinctions. Improvement is an iterative and incremental process. It involves testing and measuring the effects of small changes to optimize the reliability of services [1]. Fundamentally, improvement enhances a system by removing perceived defects and evaluating the consequences.

Innovation, by contrast, takes a less incremental approach. Innovation is about creating or adapting novel ideas that may disrupt the status quo or solve a specific problem. It is particularly useful when the end goal is vaguely defined, a new service of unclear value is developed, or the environment and underlying needs are changing rapidly [2]. Health systems face growing challenges in controlling costs while providing patient-centered care. The current model of care is being

overwhelmed by an aging population with increasingly complex and chronic diseases. Productivity in the US health care system dropped by 0.8% per year between 1990 and 2007 [3]. Even before the COVID-19 pandemic, these conditions created an environment ripe for disruption, and innovation has become an indispensable approach to solving health care's most daunting challenges.

The health care sector proved its resilience and flexibility by adapting to novel and rapidly evolving conditions as the pandemic unfolded [4]. However, as we enter the postpandemic era, health care organizations must enhance their capacity for innovation to meet pressing medical, social, and fiscal demands [4].

The Rise of Innovation Centers

Within the last 10 years, a growing number of health systems have set up innovation centers [5], with at least 110 in the United States [6,7]. Motivated by payment reforms, the potential for new revenue streams, changing patient expectations, and the prospect of losing out to competitors, health system leaders are betting that innovation centers will enable them to offer more patient-centered, affordable care. Their focus is often on creating new services that may be unrelated to current care processes. This is distinct from classic quality improvement [8] and broader than delivery science [9] since it also involves exploring new business models [10]. Business model innovation can include the design of new services, but it also considers modifying internal cost structures or revenue sources, as well as partnerships and delivery channels [11]. Since they focus on generating and eliminating novel options, and some strategies may undermine the dominant business model, innovation centers require a different set of skills and incentives to succeed within a large organization [12]. While the idea of concentrating innovation efforts in one center has been challenged [13], this model is becoming more common and is likely a feasible starting point [14]. Looking to innovation methods that have galvanized change in other industries like user-centered design, agile development, and novel digital tools, they aim to create new scalable health services [12]. Value-based payment reforms by the US federal government and other payers are creating incentives to experiment with ways to deliver care that are independent of current processes.

Innovation centers are new enough that health system leaders are still learning how to build and manage them [15]. This novelty is yielding a variety of approaches as shown by a prior Commonwealth Fund survey of innovation centers and by several case studies [6,16,17]. We examined the responses from 33 innovation centers to identify centers that appeared to have developed and spread an innovative model of care at the time, and found 10, to which we added another 3 through snowball sampling. This paper enriches the existing literature by further elucidating the link between the goals of various innovation centers and the structures built by organizations to fulfill those goals. While the internal capabilities of health systems in the discipline of service design and business model innovation are where quality improvement was two decades ago, we feel that these skills are essential to the development of sustainable, person-centered health systems.

The Aims and Organization of Innovation Centers

Nearly all centers share the aim of improving patient health status and patient experience while reducing costs. However, in practice, their strategic focus differs. Centers may focus on quality improvement, streamlining back-end processes, internal cost reduction, or revenue generation. These areas of focus suggest different structures, with examples of centers reporting to leaders in clinical quality, information technology (IT), or finance. Some emphasize improving clinical quality and operations within the organization, while others seek to create scalable models of care or new technologies with the hope of generating revenue externally. For example, the Virginia Mason Institute (hereafter referred to as Virginia Mason) focuses almost entirely on improving clinical quality within the organization (based on an adaptation of the Toyota Production System). The center is based within their quality improvement department, and the center's leadership reports to the senior vice president of quality and safety. NewYork-Presbyterian Innovation Center also focuses on internal operations, but its innovation center is distinct because it is based within the IT department and its staff report to the chief information officer. A team of IT experts and managers support system-wide implementation of mature mobile technologies to improve access to care and reduce costs. In 2019, the Hauser Institute for Health Innovation was launched to focus specifically on advancing telemedicine capabilities.

The University of Pittsburgh Medical Center (UPMC) Enterprises emphasizes the objectives of reducing costs and increasing revenue and is co-led by the treasurer of the health system. It includes a venture fund that seeds investments in start-up companies with products or services that can both support UPMC's internal operations and generate revenue from returns on investments. Northwell Ventures is primarily focused on venture investing to generate revenue. It is a distinct division within the health system that consists of an investment fund and a team of finance and management experts that work with internal and external partners to develop new technologies or models of care, but scale them up internally only if they identify interested outside investors, signaling an external market and a higher likelihood of financial return. While all of these examples aim to fulfill the broad goal of improving quality through innovation, their strategies can be delineated in practice according to their focus, structures, and key performance indicators.

Staffing and Expertise to Support an Innovation Center

Modern approaches to innovation seek a deep understanding of the user experience along with rapid prototyping and iterative testing [5]. Health systems have not traditionally employed individuals with expertise in user-centered design or entrepreneurial management methods such as lean startup. Health systems must choose between training current staff,

hiring staff with new skills, or partnering with external groups. Efforts to build internal capacity include engaging current staff through internal grants, crowdsourcing for new ideas, asking employees to vote on potential strategies to adopt, and training staff in new methods. In some cases, innovation centers create new leadership positions, such as chief innovation officer, and hire new types of staff, such as designers, developers, entrepreneurs, or finance experts. Some build facilities for rapid prototyping, offsite clinics that serve as test environments, and some have tried to build new information systems infrastructure that connects data from health systems, patients, and software and application developers [18]. Some centers hire new types of staff who work in separate divisions, while others are embedded within frontline care provision. For example, UPMC Enterprises employs engineers, designers, and clinicians who have dedicated time to develop new technologies and models of care. By contrast, Virginia Mason has employed staff to train clinical and administrative staff throughout the larger organization in methods like user-centered design while assisting them as they redesign clinical services while engaged in routine service delivery. Some groups partner externally to acquire new skills, like the Cedars-Sinai Accelerator, a collaboration between the Cedars Sinai-Medical Center and Techstars, which engaged health system leaders to identify priority areas and select start-ups to address them in cohorts with US $100,000 in financial support and mentoring to launch in a clinical environment. Others have engaged design firms and included them on advisory boards, such as the Mayo Clinic Center for Innovation engaging with the firm IDEO.

Strategic Decisions About Structuring Innovation Efforts

Many health systems displayed remarkable agility as they rapidly shifted in-person services to virtual platforms during the pandemic to reduce in-person care, but can they build this into a core capability postpandemic? There are several questions health system leaders should consider as they decide on the best structure and focus for their innovation efforts. First, is the primary goal to improve core business or develop new businesses and service lines? Second, what relevant expertise and competencies already exist in-house? These assets could include software developers, designers, technology transfer offices, or connections with early-stage entrepreneurial companies. What level of risk is the organization willing to tolerate? Those willing to tolerate higher risk may wish to go beyond improving current core activities to take on higher-risk, higher-reward options. Those with lower risk tolerance may wish to stay closer to improving core functions and focus on innovations that protect existing business from the changing landscape. What form should organizational investment in the innovation center take? Options range from smaller investments, such as partnerships and internal training of current staff, to larger investments in seed funding for a center that is expected to become financially independent, to ongoing core funding to build new facilities and hire new types of staff. Lastly, can these goals be achieved within an existing department, or do they require a new, dedicated center?

What Have Innovation Centers Achieved to Date?

There are few formal evaluations of the results of innovation center activities. As health care organizations are just beginning to learn what innovation centers can achieve, most have taken a flexible approach to evaluating them. For example, the Mayo Clinic Center for Innovation is often assigned strategy development projects, such as imagining the future of payment or quality improvement. These are tasks that other divisions are ill-equipped to pursue, and that may not have easily measurable short-term impacts. Nevertheless, there are successful examples of innovation centers improving quality, reducing costs, or generating revenue. On the quality improvement front, Virginia Mason developed and refined a clinical care protocol that reduced the time from onset of sepsis to start of treatment from 6 hours to 1 hour and refined the protocol to further reduce the time to 30 minutes, a success well beyond what other organizations have achieved. While this resembles quality improvement in approach (changes in the responsibilities of clinical staff), the magnitude of the improvement (12-fold reduction in time to treatment) is in line with what one expects from breakthrough innovations. This highlights the potential overlap between the outcomes of multiple small tests of change and radical new practices.

In the area of cost reduction, UPMC Enterprises provided seed funding to a start-up that refined natural language processing to improve coding accuracy, quadrupled coder productivity, and identified previously missed conditions to risk adjust patients for Medicare Advantage. In this instance, UPMC went from launching an untested technology developed by an external start-up to achieving system-level cost savings within one year. For revenue generation from new sources, UPMC Enterprises developed an internal consulting model for their transition to an accountable care organization. UPMC then turned the team and its protocol into an independent company called Evolent Health. From a US $38 million initial investment, the company had an initial public offering of over US $1 billion, and UPMC's stake grew to US $300 million [19,20]. While this illustrative example showed benefit in a short time frame, the centers took many years to build this capacity and have engaged in many projects before having successes like this.

How Have Innovation Centers Fared During the COVID-19 Pandemic?

Managing the response to the COVID-19 pandemic made innovation an organizational imperative, and innovation centers across the country responded in various ways. An overarching lesson from the COVID-19 response was that innovation rapidly became an organizational, rather than departmental, imperative. The task of innovating was no longer the realm of those in formal roles within innovation centers; innovation was on display as a core function for all health system leaders. Further, innovations largely focused on improving or maintaining access to services during a time of physical distancing (often through remote care), as opposed to focusing on improving quality.

UPMC Enterprises continued to invest heavily in health care innovations throughout the pandemic. UPMC Enterprises led the development of a digital patient portal for virtual appointment scheduling, expanded telemedicine capacity, and helped portfolio companies grow [21]. Butterfly Network, a medical imaging company in which UPMC Enterprises invested, went public through a US $1.5 billion acquisition deal [22]. Notably, the organization also announced a US $1 billion investment in the life sciences by 2024 [21].

Cedars-Sinai Medical Center continued to host its start-up accelerator program, albeit virtually [23]. Techstars ran a Global Startup Weekend, bringing together thousands of innovators globally to collaborate on creative solutions to the unique challenges posed by the COVID-19 pandemic [24]. The Hauser Center for Health Innovation continued to focus on innovations related to remote patient monitoring and increasing telemedicine capabilities more broadly [25].

Virginia Mason continued to improve internal quality and quickly adapted to shifting supply chain metrics and guidance from the US Centers for Disease Control and Prevention (CDC) [26]. The organization also accelerated its multiyear plan to increase telehealth capabilities [26].

Conclusion

Innovation is a set of approaches, not a destination. While the optimal forms, goals, and impact of innovation centers are still emerging, the fiscal pressures of new payment models and the potential for new digital health technologies to challenge current delivery models are real. Most health care organizations remain focused on regulatory compliance, using quality improvement to make relatively small improvements to existing processes. But this may not yield the changes needed to thrive in the future. The early experiences of innovation centers illustrate the variety of paths available to those seeking to grow their organization's capacity for innovation. As health systems adapt to the postpandemic era, it remains to be seen whether innovation centers will meet the growing medical, technological, and fiscal demands.

Acknowledgments

Lovisa Gustassfson read an early draft and provided helpful comments. This work was funded in part through a Harkness Fellowship in Policy and Practice from the Commonwealth Fund.

Conflicts of Interest

None declared.

References

1. Adapting essential care during the COVID-19 pandemic. Institute for Healthcare Improvement. URL: http://www.ihi.org/communities/blogs/africa-adapting-essential-care-during-the-covid-19-pandemic [accessed 2021-07-08]
2. Mate K. Innovation is no lightning strike: how to maximize the value of your new ideas. Institute for Healthcare Improvement. URL: http://www.ihi.org/communities/blogs/innovation-is-no-lightning-strike-how-to-maximize-the-value-of-your-new-ideas [accessed 2021-07-08]
3. Buescher B, Viguerie P. Thriving under disruption: how to succeed in the years ahead. McKinsey & Company. 2014. URL: https://healthcare.mckinsey.com/wp-content/uploads/2020/02/MCK_PayorBook_13-24_DisruptiveChange_R5.pdf [accessed 2021-07-08]
4. Cohen D, Furstenthal L, Jansen L. Healthcare innovation: building on gains made through the crisis. McKinsey & Company. URL: https://tinyurl.com/2p8d3scv [accessed 2021-07-08]
5. Pizer SD, Frakt AB, Sheetz K, Clancy C. Testing novel payment and delivery approaches through the Veterans Health Administration's new center for innovation. Ann Intern Med 2019 Jan 15;170(2):112-113. [doi: 10.7326/M18-2225] [Medline: 30583295]
6. Hostetter M, Klein S, McCarthy D, Hayes SL. Findings from a survey of health care delivery innovation centers. The Commonwealth Fund. URL: https://www.commonwealthfund.org/publications/other-publication/2015/apr/findings-survey-health-care-delivery-innovation-centers [accessed 2021-07-08]
7. Dyrda L. 66 hospitals and health systems with innovation programs. Becker's Hospital Review. URL: https://www.beckershospitalreview.com/lists/66-hospitals-and-health-systems-with-innovation-programs-2018.html [accessed 2021-07-08]
8. Bhattacharyya O, Blumenthal D, Stoddard R, Mansell L, Mossman K, Schneider EC. Redesigning care: adapting new improvement methods to achieve person-centred care. BMJ Qual Saf 2019 Mar;28(3):242-248 [FREE Full text] [doi: 10.1136/bmjqs-2018-008208] [Medline: 30242019]
9. Lieu TA, Madvig PR. Strategies for building delivery science in an integrated health care system. J Gen Intern Med 2019 Jun;34(6):1043-1047 [FREE Full text] [doi: 10.1007/s11606-018-4797-8] [Medline: 30684194]
10. Fieldston E, Terwiesch C, Altschuler S. Application of business model innovation to enhance value in health care delivery. JAMA Pediatr 2013 May;167(5):409-411. [doi: 10.1001/jamapediatrics.2013.1221] [Medline: 23549620]
11. Ostervalder A, Pigneur Y. Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers, 1st Edition. Hoboken, NJ: John Wiley and Sons; 2010.

12. Ries E. The Startup Way: How Modern Companies Use Entrepreneurial Management to Transform Culture and Drive Long-Term Growth, 1st Edition. New York, NY: Crown Publishing Group; 2017.
13. Jain SH. The health care innovation bubble. Healthc (Amst) 2017 Dec;5(4):231-232. [doi: 10.1016/j.hjdsi.2017.08.002] [Medline: 28890260]
14. Bhattacharyya O, Blumenthal D, Schneider E. Small improvements versus care redesign: can your organization juggle both? NEJM Catalyst. 2018. URL: https://catalyst.nejm.org/doi/full/10.1056/CAT.18.0282 [accessed 2021-07-08]
15. Bates DW, Sheikh A, Asch DA. Innovative environments in health care: where and how new approaches to care are succeeding. Health Aff (Millwood) 2017 Mar 01;36(3):400-407. [doi: 10.1377/hlthaff.2016.1311] [Medline: 28264940]
16. Hostetter M, Klein S, McCarthey D. Enabling change in health care. Commonwealth Fund. URL: https://medium.com/@CommonwealthFund/penn-medicine-center-for-health-care-innovation-enabling-change-612703a8f53b [accessed 2021-07-08]
17. Hughes L, McCarthy D, Audet A, Klein S. Designing more affordable and effective health care. Commonwealth Fund. URL: https://medium.com/@CommonwealthFund/designing-more-affordable-and-effective-health-care-97273e61b0a4 [accessed 2021-07-08]
18. Bresnick J. UCSF takes aim at health data interoperability with new platform. Health IT Analytics. URL: https://healthitanalytics.com/news/ucsf-takes-aim-at-health-data-interoperability-with-new-platform [accessed 2021-07-08]
19. Barkholz D. Hospital-based venture funds bet big on health startups. Modern Healthcare. URL: https://tinyurl.com/2ry2d5v9 [accessed 2021-07-08]
20. Health Catalyst becomes latest UPMC Enterprises portfolio company to go public. UPMC Enterprises. 2019. URL: https://enterprises.upmc.com/blog/health-catalyst-go-public/ [accessed 2021-07-08]
21. UPMC Enterprises 2020 year in review. UPMC Enterprises. URL: https://enterprises.upmc.com/app/uploads/2021/12/UPMC-Enterprises-2020-Year-in-Review.pdf [accessed 2021-07-08]
22. Hale C. Hand-held ultrasound developer Butterfly Network to go public through $1.5B acquisition deal. Fierce Biotech. URL: https://tinyurl.com/4km74paw [accessed 2021-07-08]
23. Cedars-Sinai Accelerator goes virtual. Cedars-Sinai. 2021 Feb 18. URL: https://www.cedars-sinai.org/newsroom/cedars-sinai-accelerator-goes-virtual/ [accessed 2021-07-08]
24. Techstars. 2020 Apr 10. URL: https://www.techstars.com/newsroom/techstars-and-differential-ventures-unite-to-fight-covid-19 [accessed 2021-07-08]
25. Drees J. 'It's about people, process and tech': 1-year post launch, NewYork-Presbyterian innovation institute leaders talk telehealth growth, future. Becker's Hospital Review. 2020 Jul 2. URL: https://tinyurl.com/yckwcb8d [accessed 2021-07-08]
26. Robeznieks A. Virginia Mason's rapid COVID-19 response had 18 years of preparation. American Medical Association. 2020 Jun 05. URL: https://www.ama-assn.org/practice-management/sustainability/virginia-mason-s-rapid-covid-19-response-had-18-years [accessed 2021-07-08]

Abbreviations

CDC: US Centers for Disease Control and Prevention
IT: information technology
UPMC: University of Pittsburgh Medical Center

Edited by G Eysenbach; submitted 30.09.21; peer-reviewed by T Lieu, A Ramjewan; comments to author 28.10.21; revised version received 19.12.21; accepted 29.12.21; published 10.02.22.

Please cite as:
Bhattacharyya O, Shapiro J, Schneider EC
Innovation Centers in Health Care Delivery Systems: Structures for Success
J Med Internet Res 2022;24(2):e33961
URL: https://www.jmir.org/2022/2/e33961
doi:10.2196/33961
PMID:35142631

Viewpoint

Passive Data Use for Ethical Digital Public Health Surveillance in a Postpandemic World

John L Kilgallon[1,2*], BA; Ishaan Ashwini Tewarie[1,3,4,5*], MD; Marike L D Broekman[1,4,5], MD, JD, PhD; Aakanksha Rana[1,6], PhD; Timothy R Smith[1], MD, MPH, PhD

[1]Computational Neurosciences Outcomes Center, Department of Neurosurgery, Brigham and Women's Hospital, Boston, MA, United States
[2]Department of General Internal Medicine, Brigham and Women's Hospital, Boston, MA, United States
[3]Faculty of Medicine, Erasmus University Rotterdam, Rotterdam, Netherlands
[4]Department of Neurosurgery, Haaglanden Medical Center, The Hague, Rotterdam, Netherlands
[5]Department of Neurosurgery, Leiden Medical Center, Leiden, Netherlands
[6]McGovern Institute for Brain Research, Massachusetts Institute of Technology, Boston, MA, United States
[*]these authors contributed equally

Corresponding Author:
John L Kilgallon, BA
Computational Neurosciences Outcomes Center
Department of Neurosurgery
Brigham and Women's Hospital
75 Francis Street
Boston, MA, 02115
United States
Phone: 1 9145395441
Email: jkilgallon@bwh.harvard.edu

Abstract

There is a fundamental need to establish the most ethical and effective way of tracking disease in the postpandemic era. The ubiquity of mobile phones is generating large amounts of passive data (collected without active user participation) that can be used as a tool for tracking disease. Although discussions of pragmatism or economic issues tend to guide public health decisions, ethical issues are the foremost public concern. Thus, officials must look to history and current moral frameworks to avoid past mistakes and ethical pitfalls. Past pandemics demonstrate that the aftermath is the most effective time to make health policy decisions. However, an ethical discussion of passive data use for digital public health surveillance has yet to be attempted, and little has been done to determine the best method to do so. Therefore, we aim to highlight four potential areas of ethical opportunity and challenge: (1) informed consent, (2) privacy, (3) equity, and (4) ownership.

(J Med Internet Res 2022;24(2):e30524) doi:10.2196/30524

KEYWORDS

passive data; public health surveillance; digital public health surveillance; pandemic response; data privacy; digital phenotyping; smartphone; mobile phone; mHealth; digital health; informed consent; data equity; data ownership

Background

In the wake of the COVID-19 pandemic, a more effective and ethically sound system for tracking disease is necessary. In recent years, due to their ubiquity, mobile phones—and the data they collect—have become a potential tool for tracking disease on a broad scale. These devices generate massive amounts of passive data, or information collected without the active participation of the user [1]. However, the current utilization of these data for public health crises is limited; such data are predominantly used for basic contact tracing via Bluetooth or GPS [2]. Emerging technologies, such as digital phenotyping, defined as moment-by-moment quantification of the individual-level human phenotype in situ using data from personal digital devices, allow for continuous monitoring of individuals' health, which has previously been impossible [3]. Attempts to employ these data for digital public health surveillance, defined as "ongoing systematic collection, analysis, and interpretation of data [not generated with the primary goal of surveillance], integrated with the timely dissemination of these data to those who can undertake effective prevention and control activities," are currently being undertaken; public health

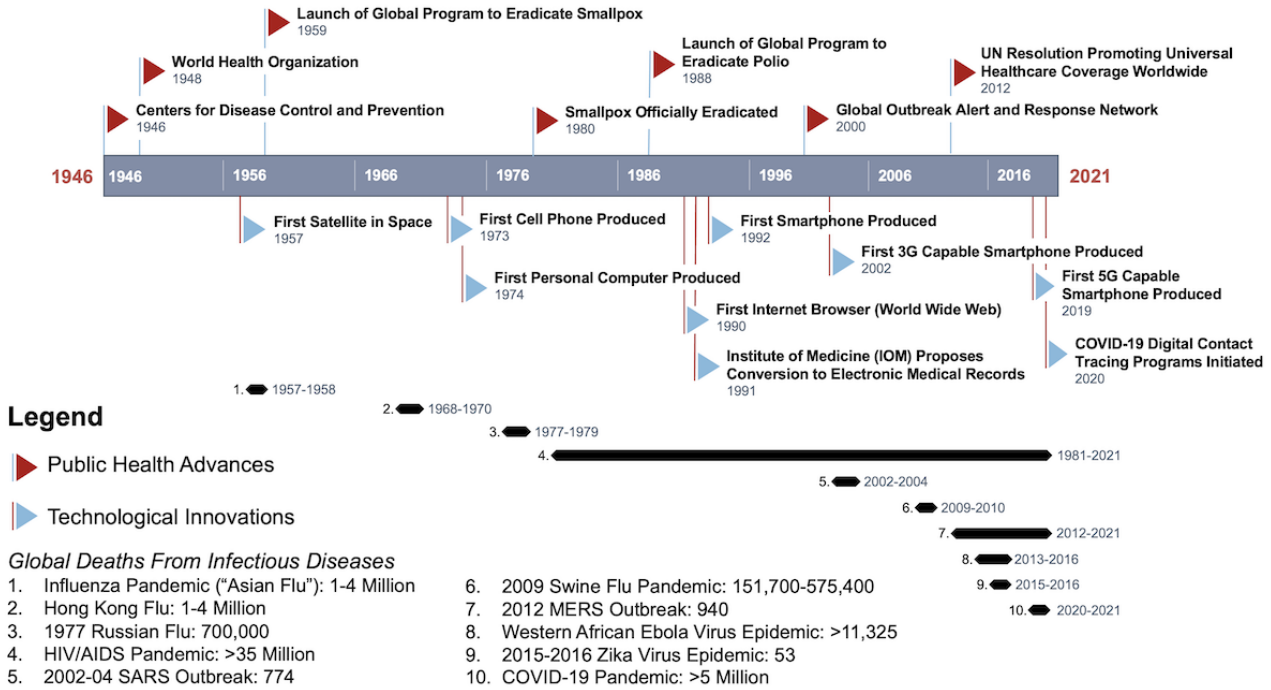
officials must look to history and current moral frameworks to avoid past mistakes and ethical pitfalls [4,5]. Thus, this viewpoint uses a scoping literature review and novel arguments to aid policy makers in critically analyzing how passive data might be used for digital public health surveillance ethically, with a particular focus on lessons to be learned from history.

A Historical Perspective on Surveillance: Proactive Versus Reactive Interventions

In 1854, John Snow laid the groundwork for modern epidemiology by disabling a cholera-contaminated water pump in Soho, London. By saving lives, Snow's experiment and subsequent health policy advances promoting hygiene demonstrated the fundamental need for proactive public health intervention in times of crisis [6]. Conversely, during the 1918 influenza pandemic, governments met the disease with deliberate ignorance [7]. Death tolls rose to 50 million worldwide, and citizens were forced to implement makeshift systems of social distancing, such as mask wearing and displaying anti-spitting signs [8,9]. The disjointed nature of the response is characteristic of the reactive approach; without unified guidelines, decided upon beforehand, citizens are left with little framework on which to base their decisions [10]. A juxtaposition can be found in the 2003 severe acute respiratory syndrome (SARS) epidemic. Although heavily affected countries developed protocols from which they have benefitted during the COVID-19 pandemic, the United States did not face the direct effects of SARS, resulting in very few steps being undertaken to prepare for future outbreaks [11-13]. Likewise, citizens in SARS-affected countries were more willing to adhere to interventions [13]. Operating with preordained protocols, these SARS-affected countries found far greater early success against COVID-19.

Pandemic response progress has fluctuated based on contemporary national politics, social norms, and the scientific understanding of diseases through history [14]. Generally, examples of health crises demonstrate that coordinated surveillance by officials and public adherence to guidelines are integral to limiting disease spread (Figure 1) [15-23]. This must be kept in mind when turning toward the future to maximize the impact of emerging technologies such as digital phenotyping for health surveillance systems. Now is the time to make policy decisions, as the choices made about passive data use for digital public health surveillance in the years following the COVID-19 pandemic have the potential to impact our lives profoundly. Although discussions of pragmatism or economic issues tend to guide digital public health surveillance decisions, ethical questions of informed consent, data privacy, data equity, and data ownership are the foremost public concerns; a consequence of not addressing these issues is eroding trust in governmental institutions and science [24]. Surveillance measures, therefore, must be functional and within ethical guidelines.

Figure 1. A timeline of modern public health advances, technological innovations, and pandemics and disease outbreaks. MERS: Middle East respiratory syndrome; SARS: severe acute respiratory syndrome: UN: United Nations.

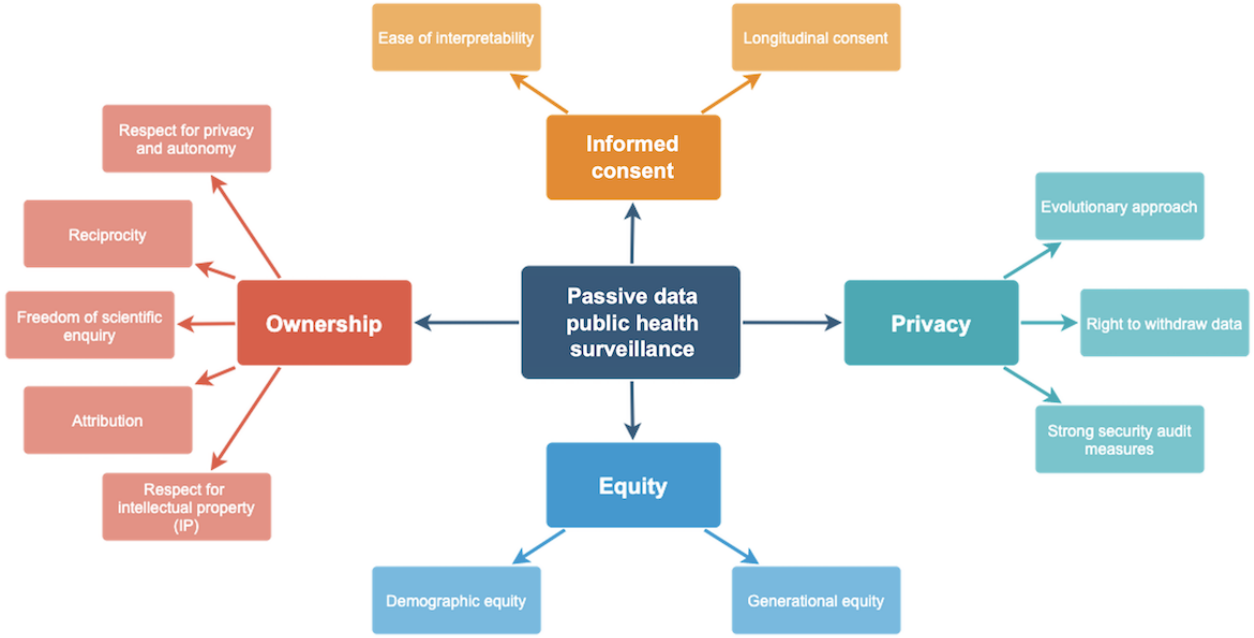


Passive Data: Pandemic Surveillance Pearls and Promises

The momentum of technological innovations has inspired a new age of digital solutions in public health. In 2014, the number of mobile phone subscriptions surpassed the number of people on the planet (~7.2 billion), illustrating a new parity between humans and devices on which policies could capitalize [3]. Early in the pandemic, governments and technology corporations deployed contact-tracing systems using Bluetooth interactions widely [2]. However, this type of solution is limited; it is reactive, only applicable after disease transmission has likely already occurred.

Recent innovations in using passive data, such as digital phenotyping, have the potential to measure a subject's physical and mental well-being and will allow for a far more accurate level of surveillance [3]. Passive data encompass various streams, including GPS, accelerometer, text, and call log data, and have been employed in a wide range of clinical settings, from monitoring spinal surgery patients' recoveries to tracking relapses in patients with schizophrenia [25-37]. The significant benefits of using passive data for digital public health surveillance include the data's objectivity, quantifiability, and continuous nature [1]. This contrasts with the use of active data, such as patient-reported outcome measures, which rely on subjective measurements that are harder to quantify accurately [1]. The sophisticated analysis of passive data for digital public health surveillance has yet to be attempted on a large scale, however, and there is no academic consensus on how best to do so [38-40]. Thus, this paper puts forth four challenges that must be addressed when considering the ethics of digital public health surveillance: (1) the level of transparency during the consent process (informed consent), (2) the anonymity and security measures taken to protect the data (privacy), (3) the equitable distribution of benefit from digital public health surveillance (equity), and (4) the determination of who has rights to the data (ownership; Figure 2) [24].

Figure 2. Goals for ethical passive data public health surveillance.



Informed Consent

Confirming that the public genuinely provides informed consent becomes challenging as information is increasingly digitized. The platform for medical surveillance has moved from controlled (eg, doctors' offices) to uncontrolled (eg, smartphones) environments. In examining the challenges of gaining patients' permission to use their data, two main pillars arise: (1) ease of interpretability and (2) longitudinal consent. Trust between officials and the public can be maintained as long as the information provided is understandable and interpretable. Levels of patient comprehension comparable to current consent procedures are attainable through mobile health consenting programs, assuming care is taken to design the interface intuitively [41]. Tools such as social annotations, live feedback, and visual aids have been suggested to help accomplish this goal [1,41,42]. However, there is little evidence to show that no current method of consenting patients achieves adequate patient comprehension [43]. Customizable, interactive, and educational consent forms to be evaluated on the patients' own time, rather than in person, would help to streamline the process and relieve stress on both sides [44,45]. Likewise, it would allow for greater regulation of bias on the part of the provider, which has been shown to play a role in patients' ability to give consent freely [46].

The perpetuity of consent has also come into question, as the technology used is inherently complex and often outside of patients' or providers' realms of expertise [47]. Therefore, the growth of digital public health surveillance could entail a move from the current practice of a "consent or anonymize" approach to one of "consent for governance" [1]. Under the former, the patient gives "broad consent" under the assumption that data will be implemented later using pseudonymization; this allows patients no long-term rights to their data. Conversely, the latter incorporates the context in which an entity hopes to use the data, making it conditional. Thus, gaining initial patient consent does not guarantee longitudinal ethical viability. A continued effort to maintain a reliable open source of information for patients is necessary.

Privacy

Experts note five distinct threats to data privacy under digital surveillance: the invisibility, inaccuracy, immortality, marketability, and identifiability of data [48]. Given this, we put forth three main pillars to be upheld in maintaining data

privacy: (1) an evolutionary approach, (2) the right to withdraw data, and (3) strong security audit measures. Authorities remark that if surveillance practices currently being conducted digitally were carried out in person (eg, a third party reading patients' text messages and following them from location to location), rather than invisibly, it would be unacceptable [48], yet this level of observation would be inherent to digital public health surveillance systems. Further, machines interpret data literally; therefore, the analysis of specific actions (eg, unknowingly dropping one's phone on the street) through passive data could lead to inaccurate conclusions (eg, a traumatic fall requiring emergency medical services) [48]. Third, researchers note that the immortality of data inevitably results in elevated risk, as even the lowest possible risk over an extended period translates to high overall risk [48]. The marketability of data is also a concern, as it is the entities that collect the data (institutions), not those that generate the data (patients) that are currently compensated, creating an incentive for unethical practices [48]. Finally, it has been demonstrated that anonymized patient data can be reidentified using machine learning [49]. Thus, the identifiability of the data could have wide-ranging consequences, from fostering attempts to shape political opinions based on one's health profile to allowing one's prognosis to have an impact on their ability to be hired [48].

Innovative multisectoral approaches will be needed going forward to prevent breaches and maintain requisite data privacy and security. The Health Insurance Portability and Accountability Act (HIPAA) in the United States and the new General Data Protection Regulation (GDPR) in the European Union were enacted to ensure legal repercussions for those who break data confidentiality. However, in the digital age, it has been posited that all data can be used as health data [48]. Likewise, as these regulations only apply to personal health information in its traditional form, they have become inadequate and outdated. Digital public health surveillance systems must continually be refined to incorporate the idiosyncrasies involved in in vivo health surveillance. Current efforts to accomplish this goal include innovations such as the cryptographic and differential privacy approach, which makes patient data less recognizable and reidentifiable, as well as federated learning, which promotes the idea of building systems without any data sharing [50]. Additionally, individuals must be allowed the right to request deletion of their health data at any time and for any reason to restrict data perpetuity. Strong privacy and security audit policies must become commonplace, as complacency with security measures could lead to areas of weakness to be exploited. Furthermore, continuous communication from the entities storing the data of its security and trustworthiness will be necessary, so as not to foster a "generalization of suspicion," where patients feel as though they are guilty until proven innocent [51].

Equity

In recent history, the unchecked or blind usage of innovative technologies has exacerbated existing health care inequalities, reflected in the form of biased data, such as the own-race bias phenomenon in facial recognition [52]. Passive data–driven solutions, then, should be thought of as tools that could reduce disparities by widening access to public health surveillance, but only if implemented mindfully, as it is unclear which subpopulations are truly at the greatest risk of these health inequities [1,53]. To ensure fairness in passive data use, a balance must be struck between two principles: (1) demographic equity and (2) generational equity. Low socioeconomic status populations and people of color are often marginalized and burdened by negative social determinants of health [54]. Passive data can potentially reduce health inequities for these populations, as conventional determinants such as lack of insurance and access to health care facilities could theoretically be circumvented [54-56]. In 2019, smartphone ownership was fairly consistent across races, with 82% of White people, 80% of Black people, and 79% of Hispanic people owning a smartphone, compared to more variable health care coverage status (92.2%, 90.4%, and 83.3%, respectively) [57,58]. Passive data use also has the potential to improve digital public health surveillance globally, as smartphone ownership in low- and middle-income countries (LMICs) continues to increase rapidly. In 2019, the median smartphone ownership was 45% in emerging economies, up from 37% in 2015 [59,60]. In LMICs, where access to adequate health care can be scarce, strengthening access to these technologies could be a viable supplement.

However, unrestrained use of technology in passive data collection and analysis could also introduce health inequities such as preventing those without technological access or with physical, age-related, disease-related, or mental impairments from receiving equitable care [61]. One population that might be left behind by increased digital public health surveillance is older adults. Although all US adults older than 65 years have access to health coverage through Medicare, only 55% owned a smartphone in 2019 [57,58]. Potential solutions to generational inequities include working with manufacturers to design digital public health surveillance technology and services keeping in mind how older populations specifically might perceive and use them, as well as implementing optimized plans and models to provide all people with affordable high-speed internet access. Overall, disparities in technological access both in the United States and globally must be tracked and actively combatted if digital public health surveillance systems are to provide equitable levels of care across demographics and generations.

Ownership

Though there is no current consensus on how best to address passive data storage and ownership issues, experts have put forth potential solutions. The majority of these involve policy interventions to restrict single institutions from monopolizing databases [39]. For example, some suggest public policy to create networks of patients with data sharing responsibility. In contrast, others promote policies that consider the rights to one's health data to be civil liberty [39,62,63]. Another possible solution that has gained traction is a paradigm shift in governing data access from ownership to custodianship [64]. This aligns with the theory that big data cannot be "owned" in a traditional sense, but instead should be guarded and overseen [64]. For this new structure to be viable, five principles must be upheld: (1) respect for privacy and autonomy, (2) reciprocity, (3) freedom

of scientific enquiry, (4) attribution, and (5) respect for intellectual property (IP) [64]. Under this system, a balance must be struck between confidentiality and accountability, as institutions (or custodians) of the data must keep the identities of the subjects (or donors) private while also remaining beholden to them. Likewise, custodians must also be forthcoming with their findings and metadata, just as donors are with the original passive health data. Health data should be used solely for the common good, as its value makes it a target to be bought and sold by bad actors [64]. Lastly, proper credit and respect for IP will reduce the restriction of access to databases by ensuring that the sacrifices made by both donor and custodian are appreciated. These data-sharing agreements are made fittingly [64]. Pioneering policy ideas such as these will be necessary if passive data is to be implemented into digital public health surveillance successfully.

Even under the assumption that future frameworks should widen access to these data, some researchers argue for a centralized database, while others support decentralization [39,58]. The latter group argues that patients have little idea of their data's actual value. Systems such as blockchain (ie, a digital record of transactions validated by a peer-to-peer network) could serve to decentralize and better quantify the worth of individuals' health data [39,65]. Each proposal has positives and negatives. For example, centralization would increase ease of access but could leave databases more susceptible to large-scale hacks. In addition, although decentralization would allow patients to earn tangible rewards for their data, this could lead to undue influences playing a role in the decision to share one's data [63,65]. Coming to a consensus on these issues is one of the most critical next steps for advancing passive data use for digital public health surveillance.

Conclusion: Planning Now for the Future

The time to plan and prepare for the next pandemic is now. This means coming to terms with the massive potential (for better or worse) of passive data derived from personal digital devices. The potential for success of digital public health surveillance relies on the ethical viability of its implementation. Before the next pandemic, officials must ensure that we are prepared by overhauling current consenting protocols. During the next pandemic, policies must be enacted to account for all possible inequities, while maintaining trust between citizens and institutions. In the wake of the next pandemic, the long-term security of health data must be guaranteed with an assurance that it will not be used for any undue gain. These preparations must be undertaken now, proactively, using the lessons fresh in citizens' collective consciousness. Passive data and methods such as digital phenotyping can serve as the foundation upon which this improved system can be built. However, if done without historical, ethical, and practical considerations, we will be left with even more challenges than we face today.

Authors' Contributions

JLK, IAT, and AR designed the research. JLK and IAT cowrote the paper. MLDB, AR, and TRS jointly supervised this work. MLDB, AR, and TRS critically revised the paper.

Conflicts of Interest

None declared.

References

1. Maher NA, Senders JT, Hulsbergen AFC, Lamba N, Parker M, Onnela J, et al. Passive data collection and use in healthcare: A systematic review of ethical issues. Int J Med Inform 2019 Sep;129:242-247. [doi: 10.1016/j.ijmedinf.2019.06.015] [Medline: 31445262]
2. Zastrow M. Coronavirus contact-tracing apps: can they slow the spread of COVID-19? Nature 2020 May 19. [doi: 10.1038/d41586-020-01514-2] [Medline: 32433633]
3. Onnela J, Rauch SL. Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. Neuropsychopharmacology 2016 Jun;41(7):1691-1696. [doi: 10.1038/npp.2016.7] [Medline: 26818126]
4. Norbury A, Liu SH, Campaña-Montes JJ, Romero-Medrano L, Barrigón ML, Smith E, MEmind Study Group, et al. Social media and smartphone app use predicts maintenance of physical activity during Covid-19 enforced isolation in psychiatric outpatients. Mol Psychiatry 2020 Dec 14:3920-3930 [FREE Full text] [doi: 10.1038/s41380-020-00963-5] [Medline: 33318619]
5. Shakeri Hossein Abad Z, Kline A, Sultana M, Noaeen M, Nurmambetova E, Lucini F, et al. Digital public health surveillance: a systematic scoping review. NPJ Digit Med 2021 Mar 03;4(1):41 [FREE Full text] [doi: 10.1038/s41746-021-00407-6] [Medline: 33658681]
6. Ruths MB. The lesson of john snow and the broad street pump. Virtual Mentor 2009 Jun 01;11(6):470-472 [FREE Full text] [doi: 10.1001/virtualmentor.2009.11.6.mhst1-0906] [Medline: 23199359]
7. Martini M, Gazzaniga V, Bragazzi NL, Barberis I. The Spanish Influenza Pandemic: a lesson from history 100 years after 1918. J Prev Med Hyg 2019 Mar;60(1):E64-E67 [FREE Full text] [doi: 10.15167/2421-4248/jpmh2019.60.1.1205] [Medline: 31041413]
8. Centers for Disease Control and Prevention; 2018. History of 1918 Flu Pandemic. URL: https://www.cdc.gov/flu/pandemic-resources/1918-commemoration/1918-pandemic-history.htm [accessed 2020-07-17]

9. Little B. HISTORY. 'Mask Slackers' and 'Deadly' Spit: The 1918 Flu Campaigns to Shame People Into Following New Rules URL: https://www.history.com/news/1918-pandemic-public-health-campaigns [accessed 2022-02-04]
10. McCullough M. Ethical decision making in a pandemic: where are the voices of vulnerable people? BMJ 2020 Jun 19;369:m2406. [doi: 10.1136/bmj.m2406] [Medline: 32561520]
11. Broder J. The SARS Epidemic: The American Response; Aggressive Steps, and Luck, Help U.S. Avoid SARS Brunt. The New York Times. 2003 May 05. URL: https://www.nytimes.com/2003/05/05/us/sars-epidemic-american-response-aggressive-steps-luck-help-us-avoid-sars-brunt.html [accessed 2022-02-04]
12. Person B, Sy F, Holton K, Govert B, Liang A, National Center for Inectious Diseases/SARS Community Outreach Team. Fear and stigma: the epidemic within the SARS outbreak. Emerg Infect Dis 2004 Feb;10(2):358-363 [FREE Full text] [doi: 10.3201/eid1002.030750] [Medline: 15030713]
13. Tan KYL. Singapore's Regulatory Response to COVID-19. The Regulatory Review. 2020 Jul 01. URL: https://www.theregreview.org/2020/06/15/tan-singapore-regulatory-response-covid-19/ [accessed 2021-04-15]
14. Institute of Medicine (US). Ethical and Legal Considerations in Mitigating Pandemic Disease. In: National Academies Press (US). 2007 Presented at: Institute of Medicine (US) Forum on Microbial Threats; September 19-20, 2006; Washington, DC. [doi: 10.17226/11917]
15. European Centre for Disease Prevention and Control. Geographical distribution of confirmed MERS-CoV cases by country of infection and year. European Union. URL: https://www.ecdc.europa.eu/en/publications-data/geographical-distribution-confirmed-mers-cov-cases-country-infection-and-year [accessed 2021-04-25]
16. Dawood FS, Iuliano AD, Reed C, Meltzer MI, Shay DK, Cheng P, et al. Estimated global mortality associated with the first 12 months of 2009 pandemic influenza A H1N1 virus circulation: a modelling study. Lancet Infect Dis 2012 Sep;12(9):687-695. [doi: 10.1016/S1473-3099(12)70121-4] [Medline: 22738893]
17. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis 2020 Feb 19:533-534. [doi: 10.1016/S1473-3099(20)30120-1] [Medline: 32087114]
18. Michaelis M, Doerr HW, Cinatl J. Novel swine-origin influenza A virus in humans: another pandemic knocking at the door. Med Microbiol Immunol 2009 Aug;198(3):175-183. [doi: 10.1007/s00430-009-0118-5] [Medline: 19543913]
19. Pandemic Influenza Risk Management WHO Interim Guidance. World Health Organization. 2013. URL: https://www.who.int/influenza/preparedness/pandemic/GIP_PandemicInfluenzaRiskManagementInterimGuidance_Jun2013.pdf?ua=1 [accessed 2021-04-25]
20. Summary of probable SARS cases with onset of illness from 1 November 2002 to 31 July 2003. World Health Organization. 2015. URL: https://www.who.int/publications/m/item/summary-of-probable-sars-cases-with-onset-of-illness-from-1-november-2002-to-31-july-2003 [accessed 2021-04-25]
21. 2015-16 Zika Virus Epidemic. Worldwide Outbreak. URL: https://www.worldwideoutbreak.com/blog/cool_timeline/2015-16-zika-virus-epidemic/ [accessed 2021-04-25]
22. 2014-2016 Ebola Outbreak in West Africa. Centers for Disease Control and Prevention. 2019. URL: https://www.cdc.gov/vhf/ebola/history/2014-2016-outbreak/index.html [accessed 2021-04-25]
23. HIV Surveillance Report, 2018 (Updated). Centers for Disease Control and Prevention. URL: http://www.cdc.gov/hiv/library/reports/hiv-surveillance.html [accessed 2021-04-15]
24. Simon SR, Evans JS, Benjamin A, Delano D, Bates DW. Patients' attitudes toward electronic health information exchange: qualitative study. J Med Internet Res 2009;11(3):e30 [FREE Full text] [doi: 10.2196/jmir.1164] [Medline: 19674960]
25. Aledavood T, Torous J, Triana Hoyos AM, Naslund JA, Onnela J, Keshavan M. Smartphone-based tracking of sleep in depression, anxiety, and psychotic disorders. Curr Psychiatry Rep 2019 Jun 04;21(7):49 [FREE Full text] [doi: 10.1007/s11920-019-1043-y] [Medline: 31161412]
26. Barnett I, Onnela J. Inferring mobility measures from GPS traces with missing data. Biostatistics 2020 Apr 01;21(2):e98-e112. [doi: 10.1093/biostatistics/kxy059] [Medline: 30371736]
27. Barnett I, Torous J, Staples P, Keshavan M, Onnela J. Beyond smartphones and sensors: choosing appropriate statistical methods for the analysis of longitudinal data. J Am Med Inform Assoc 2018 Dec 01;25(12):1669-1674. [doi: 10.1093/jamia/ocy121] [Medline: 30272176]
28. Barnett I, Torous J, Staples P, Sandoval L, Keshavan M, Onnela J. Relapse prediction in schizophrenia through digital phenotyping: a pilot study. Neuropsychopharmacology 2018 Jul;43(8):1660-1666. [doi: 10.1038/s41386-018-0030-z] [Medline: 29511333]
29. Berry JD, Paganoni S, Carlson K, Burke K, Weber H, Staples P, et al. Design and results of a smartphone-based digital phenotyping study to quantify ALS progression. Ann Clin Transl Neurol 2019 May;6(5):873-881 [FREE Full text] [doi: 10.1002/acn3.770] [Medline: 31139685]
30. Beukenhorst AL, Collins E, Burke KM, Rahman SM, Clapp M, Konanki SC, et al. Smartphone data during the COVID-19 pandemic can quantify behavioral changes in people with ALS. Muscle Nerve 2021 Feb;63(2):258-262 [FREE Full text] [doi: 10.1002/mus.27110] [Medline: 33118628]
31. Cote DJ, Barnett I, Onnela J, Smith TR. Digital phenotyping in patients with spine disease: A novel approach to quantifying mobility and quality of life. World Neurosurgery 2019 Jun;126:e241-e249. [doi: 10.1016/j.wneu.2019.01.297]

32. Panda N, Solsky I, Huang EJ, Lipsitz S, Pradarelli JC, Delisle M, et al. Using smartphones to capture novel recovery metrics after cancer surgery. JAMA Surg 2020 Feb 01;155(2):123-129 [FREE Full text] [doi: 10.1001/jamasurg.2019.4702] [Medline: 31657854]
33. Staples P, Torous J, Barnett I, Carlson K, Sandoval L, Keshavan M, et al. A comparison of passive and active estimates of sleep in a cohort with schizophrenia. NPJ Schizophr 2017 Oct 16;3(1):37 [FREE Full text] [doi: 10.1038/s41537-017-0038-0] [Medline: 29038553]
34. Torous J, Kiang MV, Lorme J, Onnela J. New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone Research. JMIR Ment Health 2016;3(2):e16 [FREE Full text] [doi: 10.2196/mental.5165] [Medline: 27150677]
35. Torous J, Onnela J, Keshavan M. New dimensions and new tools to realize the potential of RDoC: digital phenotyping via smartphones and connected devices. Transl Psychiatry 2017 Dec 07;7(3):e1053 [FREE Full text] [doi: 10.1038/tp.2017.25] [Medline: 28267146]
36. Torous J, Staples P, Barnett I, Sandoval LR, Keshavan M, Onnela J. Characterizing the clinical relevance of digital phenotyping data quality with applications to a cohort with schizophrenia. NPJ Digit Med 2018;1:15 [FREE Full text] [doi: 10.1038/s41746-018-0022-8] [Medline: 31304300]
37. Torous J, Staples P, Onnela J. Realizing the potential of mobile mental health: new methods for new data in psychiatry. Curr Psychiatry Rep 2015 Aug;17(8):602. [doi: 10.1007/s11920-015-0602-0] [Medline: 26073363]
38. Hand DJ. Aspects of data ethics in a changing world: where are we now? Big Data 2018 Sep 01;6(3):176-190 [FREE Full text] [doi: 10.1089/big.2018.0083] [Medline: 30283727]
39. Mirchev M, Mircheva I, Kerekovska A. The academic viewpoint on patient data ownership in the context of big data: scoping review. J Med Internet Res 2020 Aug 18;22(8):e22214. [doi: 10.2196/22214] [Medline: 32808934]
40. Mittelstadt BD, Floridi L. The ethics of big data: current and foreseeable issues in biomedical contexts. Sci Eng Ethics 2016 Apr;22(2):303-341. [doi: 10.1007/s11948-015-9652-2] [Medline: 26002496]
41. Iwaya LH, Li J, Fischer-Hübner S, Åhlfeldt R, Martucci LA. E-consent for data privacy: consent management for mobile health technologies in public health surveys and disease surveillance. Stud Health Technol Inform 2019 Aug 21;264:1223-1227. [doi: 10.3233/SHTI190421] [Medline: 31438120]
42. Chen C, Lee P, Pain KJ, Delgado D, Cole CL, Campion TR. Replacing paper informed consent with electronic informed consent for research in academic medical centers: a scoping review. AMIA Jt Summits Transl Sci Proc 2020;2020:80-88 [FREE Full text] [Medline: 32477626]
43. Agozzino E, Borrelli S, Cancellieri M, Carfora FM, Di Lorenzo T, Attena F. Does written informed consent adequately inform surgical patients? A cross sectional study. BMC Med Ethics 2019 Jan 07;20(1):1 [FREE Full text] [doi: 10.1186/s12910-018-0340-z] [Medline: 30616673]
44. Harle CA, Golembiewski EH, Rahmanian KP, Krieger JL, Hagmajer D, Mainous AG, et al. Patient preferences toward an interactive e-consent application for research using electronic health records. J Am Med Inform Assoc 2018 Mar 01;25(3):360-368 [FREE Full text] [doi: 10.1093/jamia/ocx145] [Medline: 29272408]
45. Skelton E, Drey N, Rutherford M, Ayers S, Malamateniou C. Electronic consenting for conducting research remotely: A review of current practice and key recommendations for using e-consenting. Int J Med Inform 2020 Nov;143:104271 [FREE Full text] [doi: 10.1016/j.ijmedinf.2020.104271] [Medline: 32979650]
46. Hall DE, Prochazka AV, Fink AS. Informed consent for clinical treatment. CMAJ 2012 Mar 20;184(5):533-540 [FREE Full text] [doi: 10.1503/cmaj.112120] [Medline: 22392947]
47. Helgesson G, Ludvigsson J, Gustafsson Stolt U. How to handle informed consent in longitudinal studies when participants have a limited understanding of the study. J Med Ethics 2005 Nov;31(11):670-673 [FREE Full text] [doi: 10.1136/jme.2004.009274] [Medline: 16269567]
48. Grande D, Luna Marti X, Feuerstein-Simon R, Merchant RM, Asch DA, Lewson A, et al. Health policy and privacy challenges associated with digital technology. JAMA Netw Open 2020 Jul 01;3(7):e208285 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.8285] [Medline: 32644138]
49. Na L, Yang C, Lo C, Zhao F, Fukuoka Y, Aswani A. Feasibility of reidentifying individuals in large national physical activity data sets from which protected health information has been removed with use of machine learning. JAMA Netw Open 2018 Dec 07;1(8):e186040 [FREE Full text] [doi: 10.1001/jamanetworkopen.2018.6040] [Medline: 30646312]
50. Kaissis GA, Makowski MR, Rückert D, Braren RF. Secure, privacy-preserving and federated machine learning in medical imaging. Nat Mach Intell 2020 Jun 08;2(6):305-311. [doi: 10.1038/s42256-020-0186-1]
51. Samerski S. Individuals on alert: digital epidemiology and the individualization of surveillance. Life Sci Soc Policy 2018 Jun 14;14(1):13 [FREE Full text] [doi: 10.1186/s40504-018-0076-z] [Medline: 29900518]
52. Wong HK, Stephen ID, Keeble DRT. The own-race bias for face recognition in a multiracial society. Front Psychol 2020;11:208 [FREE Full text] [doi: 10.3389/fpsyg.2020.00208] [Medline: 32210861]
53. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019 Oct 25;366(6464):447-453. [doi: 10.1126/science.aax2342] [Medline: 31649194]
54. Braveman P, Gottlieb L. The social determinants of health: it's time to consider the causes of the causes. Public Health Rep 2014;129 Suppl 2:19-31 [FREE Full text] [doi: 10.1177/00333549141291S206] [Medline: 24385661]

55. Armstrong K, Ravenell KL, McMurphy S, Putt M. Racial/ethnic differences in physician distrust in the United States. Am J Public Health 2007 Jul;97(7):1283-1289. [doi: 10.2105/AJPH.2005.080762] [Medline: 17538069]
56. Scharff DP, Mathews KJ, Jackson P, Hoffsuemmer J, Martin E, Edwards D. More than Tuskegee: understanding mistrust about research participation. J Health Care Poor Underserved 2010 Aug;21(3):879-897 [FREE Full text] [doi: 10.1353/hpu.0.0323] [Medline: 20693733]
57. Mobile Fact Sheet. Pew Research Center. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2021-04-15]
58. Keisler-Starkey K, Bunch L. Health Insurance Coverage in the United States. Washington, DC: U.S. Census Bureau; 2021 Sep 14. URL: https://www.census.gov/library/publications/2020/demo/p60-271.html [accessed 2022-02-04]
59. Poushter J. Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies. Pew Research Center. 2016. URL: https://www.pewresearch.org/global/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/ [accessed 2021-04-15]
60. Silver L. Smartphone Ownership Is Growing Rapidly Around the World, but Not Always Equally. Pew Research Center. 2019. URL: https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/ [accessed 2021-04-15]
61. Eruchalu CN, Pichardo MS, Bharadwaj M, Rodriguez CB, Rodriguez JA, Bergmark RW, et al. The expanding digital divide: digital health access inequities during the COVID-19 pandemic in New York City. J Urban Health 2021 Apr;98(2):183-186 [FREE Full text] [doi: 10.1007/s11524-020-00508-9] [Medline: 33471281]
62. Heitmueller A, Henderson S, Warburton W, Elmagarmid A, Pentland AS, Darzi A. Developing public policy to advance the use of big data in health care. Health Aff (Millwood) 2014 Sep;33(9):1523-1530. [doi: 10.1377/hlthaff.2014.0771] [Medline: 25201656]
63. Kish LJ, Topol EJ. Unpatients-why patients should own their medical data. Nat Biotechnol 2015 Sep;33(9):921-924. [doi: 10.1038/nbt.3340] [Medline: 26348958]
64. Ballantyne A. How should we think about clinical data ownership? J Med Ethics 2020 May;46(5):289-294. [doi: 10.1136/medethics-2018-105340] [Medline: 31911499]
65. Mamoshina P, Ojomoko L, Yanovich Y, Ostrovski A, Botezatu A, Prikhodko P, et al. Converging blockchain and next-generation artificial intelligence technologies to decentralize and accelerate biomedical research and healthcare. Oncotarget 2018 Jan 19;9(5):5665-5690 [FREE Full text] [doi: 10.18632/oncotarget.22345] [Medline: 29464026]

Abbreviations

GDPR: General Data Protection Regulation
HIPAA: Health Insurance Portability and Accountability Act
IP: intellectual property
LMIC: low- and middle-income country
SARS: severe acute respiratory syndrome

Edited by A Mavragani; submitted 18.05.21; peer-reviewed by Z Shakeri, H Jarke, I Mircheva; comments to author 20.07.21; revised version received 14.09.21; accepted 30.11.21; published 15.02.22.

Please cite as:
Kilgallon JL, Tewarie IA, Broekman MLD, Rana A, Smith TR
Passive Data Use for Ethical Digital Public Health Surveillance in a Postpandemic World
J Med Internet Res 2022;24(2):e30524
URL: https://www.jmir.org/2022/2/e30524
doi:10.2196/30524
PMID:35166676

Viewpoint

Building a Precision Medicine Delivery Platform for Clinics: The University of California, San Francisco, BRIDGE Experience

Riley Bove[1*], MD; Erica Schleimer[1*], BA; Paul Sukhanov[1], MSE; Michael Gilson[1], MSc; Sindy M Law[1], MS; Andrew Barnecut[1], BA; Bruce L Miller[1], MD; Stephen L Hauser[1], MD; Stephan J Sanders[1*], PhD; Katherine P Rankin[1*], PhD

UCSF Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, CA, United States

[*]these authors contributed equally

Corresponding Author:
Riley Bove, MD
UCSF Weill Institute for Neurosciences
University of California, San Francisco
1651 4th Street
San Francisco, CA, 94158
United States
Phone: 1 415 353 2069
Fax: 1 415 353 2633
Email: Riley.bove@ucsf.edu

Abstract

Despite an ever-expanding number of analytics with the potential to impact clinical care, the field currently lacks point-of-care technological tools that allow clinicians to efficiently select disease-relevant data about their patients, algorithmically derive clinical indices (eg, risk scores), and view these data in straightforward graphical formats to inform real-time clinical decisions. Thus far, solutions to this problem have relied on either bottom-up approaches that are limited to a single clinic or generic top-down approaches that do not address clinical users' specific setting-relevant or disease-relevant needs. As a road map for developing similar platforms, we describe our experience with building a custom but institution-wide platform that enables economies of time, cost, and expertise. The BRIDGE platform was designed to be modular and scalable and was customized to data types relevant to given clinical contexts within a major university medical center. The development process occurred by using a series of human-centered design phases with extensive, consistent stakeholder input. This institution-wide approach yielded a unified, carefully regulated, cross-specialty clinical research platform that can be launched during a patient's electronic health record encounter. The platform pulls clinical data from the electronic health record (Epic; Epic Systems) as well as other clinical and research sources in real time; analyzes the combined data to derive clinical indices; and displays them in simple, clinician-designed visual formats specific to each disorder and clinic. By integrating an application into the clinical workflow and allowing clinicians to access data sources that would otherwise be cumbersome to assemble, view, and manipulate, institution-wide platforms represent an alternative approach to achieving the vision of true personalized medicine.

(J Med Internet Res 2022;24(2):e34560) doi:10.2196/34560

KEYWORDS

precision medicine; clinical implementation; in silico trials; clinical dashboard; precision; implementation; dashboard; design; experience; analytic; tool; analysis; decision-making; real time; platform; human-centered design

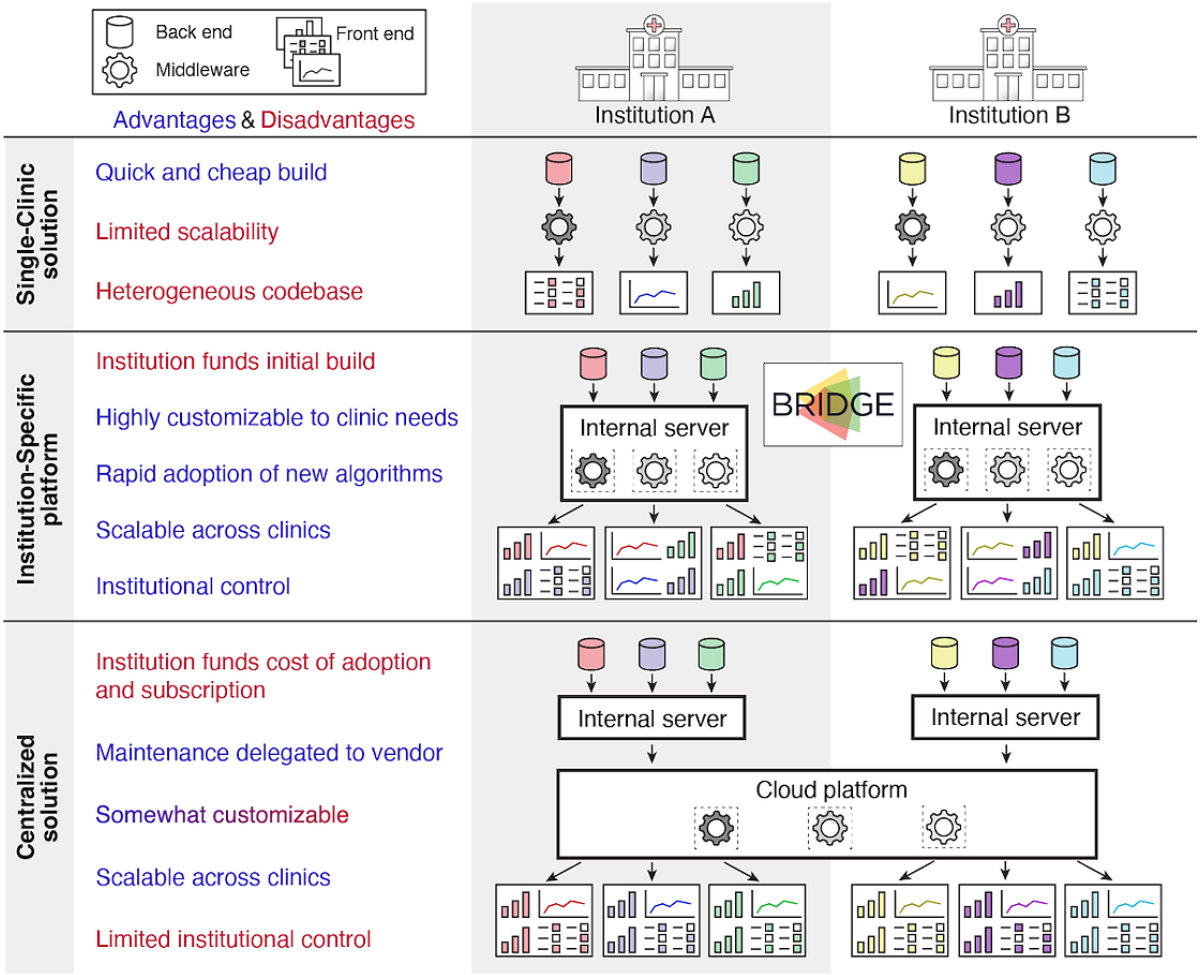
Introduction

Precision medicine holds the potential to revolutionize medicine [1-3], just as prior technological advances, such as microscopy, molecular diagnostics, and imaging, have done in the past. In the research realm, big data and artificial intelligence have yielded substantial advances that showcase the potential of precision medicine [4,5]. However, translating these advances into the clinical realm remains a challenge [6,7]. A patient is more likely to interact with complex algorithms informed by big data in the waiting room (ie, algorithms in the form of internet searches, travel directions, or tailored social media) than in the actual clinic. The medical field needs similarly intuitive interfaces that can collate the necessary patient-related data to highlight salient knowledge, pinpoint a patient's condition, predict optimal therapy, or estimate the risk of disease or death [3]. Much of the required physical infrastructure is already in place, with computers being available in most clinics and the majority of clinical data being stored in electronic health

records (EHRs). A small minority of wealthier clinics and health care systems have built custom, domain-specific interfaces into their EHRs to deliver the more complex precision medicine algorithms and visualizations that their physicians need; however, in the majority of health systems, only the most basic algorithms (eg, those for calculating BMI) are built into the EHR, while other, more sophisticated clinical indices (eg, atrial fibrillation stroke risk [8,9]) are calculated via manual entry into a public website [10].

The task of translating innovative precision medicine tools from research projects to clinical care is inhibited by a catch-22 problem. To justify the expense of building the costly computational infrastructure required to run complex algorithms on patient data, the algorithms or visualizations need to demonstrate real-world value. However, to evaluate and prove these algorithms' value, the needed infrastructure must already be in place. One solution to this conundrum is building boutique, single-clinic solutions consisting of carefully designed, specialized algorithms or data displays built within or alongside the EHR [11,12]. Although this bottom-up approach is limited in scope to a single clinical domain and thus can be comparatively quick and cost-effective to implement, scalability and rapid obsolescence are major concerns. To adapt data displays to other clinics, an institution has to maintain, secure, and update an ever-expanding heterogeneous code base across those clinics. Yet, the originating "owners" of these algorithms are often clinical researchers and physicians without the backing of an enterprise-level developer team that is equipped to manage the software as a service over several years of use (Figure 1). The opposite extreme is commercial vendors building generalized health care software suites that run on cloud-based infrastructures. Such centralized solutions address the scalability challenges of bottom-up approaches, but the emerging health system–wide products are typically far too generic to meet the medically heterogeneous and shifting requirements of individual clinics. Furthermore, adopting such solutions requires substantial institutional investment, and becoming locked into a single vendor in a rapidly evolving marketplace poses a risk.

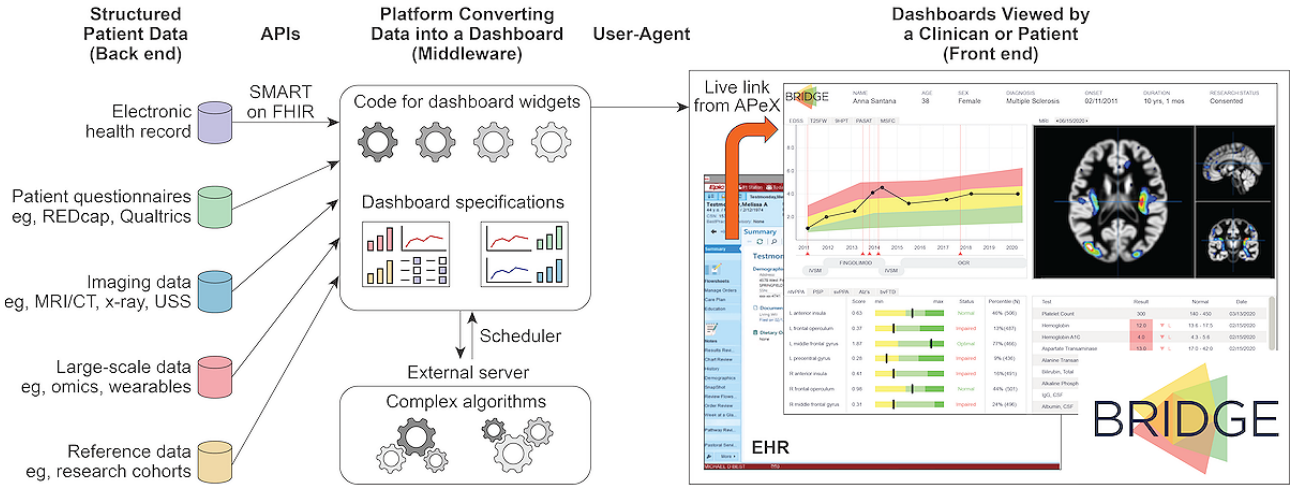
Figure 1. Approaches to delivering precision medicine results to the clinic. This figure compares the platform design elements across the following three main approaches to building clinical systems to support precision medicine: (1) single-clinic solutions, (2) institution-specific platforms like BRIDGE, and (3) centralized solutions purchased from external vendors. Key advantages (blue) and disadvantages (red) of each approach are listed.



Between these two extremes exists a third solution that solves many of the aforementioned problems. Institution-wide platforms permit rapid innovation in parallel across multiple clinics but are built on a single secure, stable, and cost-efficient

technological foundation. These platforms benefit from a common architecture built within an institutional firewall with real-time EHR access and application programming interfaces (APIs) to major (eg, REDCap [Research Electronic Data Capture; Vanderbilt University] and Radiology PACS [Picture Archiving and Communication System]) and custom data resources, which facilitate the integration of multimodal research data across all specialties. Yet, these platforms also incorporate clinic-specific visualization tools that allow clinicians to tailor the display of information. Therefore, specific research discoveries can be rapidly translated into clinical tools that fit each specialty (Figure 2). This approach strikes a balance between the fast development and flexibility of single-clinic solutions and the scalability and sustainability of centralized health care solutions while optimizing transparent institutional oversight.

Figure 2. Overview of technological components for the integrated delivery of precision medicine through an institution-specific platform like BRIDGE. The flow of information is depicted as it moves from back-end data sources, through the integrating middleware layer of light and heavy computational resources, and to the multiple, functionally distinct widgets shown in a user-facing front-end dashboard designed to reflect the needs of a single clinic. APeX: Advanced Patient-Centered Excellence; API: application programming interface; CT: computerized tomography; EHR: electronic health record; FHIR: Fast Healthcare Interoperability Resources; MRI: magnetic resonance imaging; REDCap: Research Electronic Data Capture; SMART: Substitutable Medical Applications and Reusable Technologies; USS: ultrasound sonography.



The BRIDGE platform at University of California, San Francisco (UCSF), is one example of this approach. Based on our experience with developing BRIDGE, we describe key considerations and practical steps for implementing institution-wide solutions in this rapidly progressing field to provide a road map for other health care systems considering a similar approach. We also consider future developments that will enable the medical community to quickly and comprehensively realize the potential of computational medicine to improve the lives of patients.

Consideration 1: Human-Centered Design

Overview of the Human-Centered Design of Precision Medicine Tools

For a precision medicine tool to be adopted in a clinic, it needs to provide pertinent, actionable information in a format that is appropriate to the user (either a clinician or a patient). Therefore, perhaps the most essential components of effective precision medicine tool deployment are the principles and phases of human-centered design (HCD) [13-15]. For tools targeted at medical professionals, clinician users, who are well informed, should be at the center of decisions about which technological format is the most appropriate for their workflow, which innovations in their specialty are scientifically ready for deployment at clinics, and how evaluations of tool effectiveness should be conducted to justify the continued use of such tools (Textbox 1). Many of these decisions reflect the dimensions of precision medicine, as articulated in a recent scoping review [7].

Textbox 1. Key decisions in designing a digital application for clinical research.

Key questions

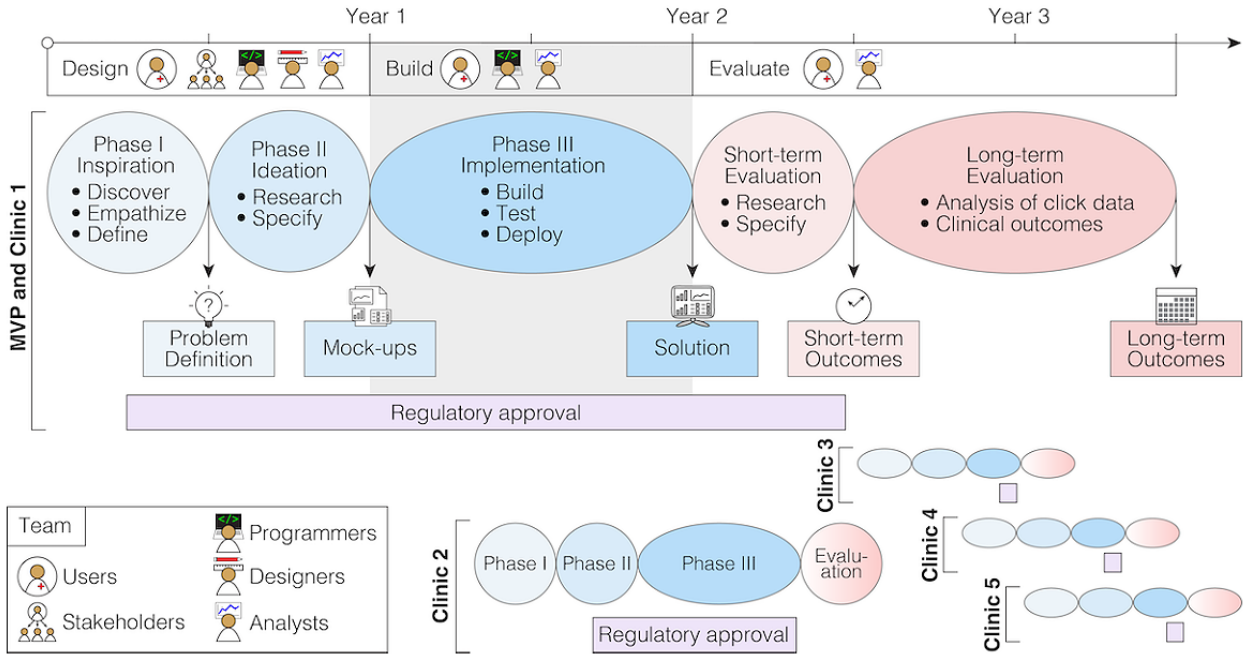
1. Who are the users (eg, clinicians, patients, and specialists)?
2. What do the users need (eg, novel data sources, novel algorithms, novel visualization, and data collation)?
3. How will it improve care (eg, patient experience, clinic efficiency, morbidity, and mortality)?
4. How does the user access the application (eg, individual log-in and authorization via an existing clinical system)?
5. Where is it hosted (eg, local server, cloud-based server, or external vendor)?
6. What is the maintenance schedule (eg, 9 AM to 5 PM on Monday to Friday or 24 hours per day year-round)?
7. What are the constraints of the system? For example, will it not write to the electronic health record? Are data behind an institutional firewall?

Practical Considerations From the BRIDGE Experience

From its inception, BRIDGE exemplified both the principles and phases of HCD [13,14]. It was conceptualized and designed according to the requirements of clinician scientists, including the project's principal investigators (manuscript authors RB, KPR, and SJS). Further, the key architectural decisions (Textbox 1) were made by applying HCD principles to engage clinician, patient, scientific, programming, design, industry, and institutional stakeholders.

The three HCD phases are also being deployed in the iterative process of adapting BRIDGE to each new clinic that is interested in a BRIDGE dashboard (Figure 3). In the "Inspiration" phase, the BRIDGE clinician scientists and programmers identify and meet with a small number of clinician champions to collaboratively define the problems to be solved to improve care in the clinician champions' clinic. They also generate ideal use cases based on that clinic's workflow, such as those for data types, data sources, and visualizations. In the "Ideation" phase, a design mock-up is shared with a broader set of intended stakeholders from that clinic to obtain their input, after which the final set of minimum viable product (MVP) specifications is derived for the dashboard, and programming begins for the jointly approved design mock-up. The finalized MVP is built in the "Implementation" phase, during which early testing is conducted by a small superuser group of clinicians who generate feedback about bugs and minor refinements. These clinic domain experts are the primary drivers for designing and conducting formal evaluations of their precision medicine tools, which include clinician users' feedback about dashboard ease, utility, and fidelity; patients' satisfaction with care; impacts on workflow, including automated click tracking; and longer-term analyses of the clinical impact, value, and cost-effectiveness of these tools. Clinical validation, technological or therapeutic innovation, or user demand may result in further cycles of design.

Figure 3. Example timelines and milestones of clinical application development. The design and development of an institution-specific platform for clinical applications, such as BRIDGE, is a multiyear undertaking (eg, 2 years for the MVP and clinic 1). Following the principles and phases of human-centered design ensures the development of a product that meets the needs of the users and the requirements of the institution. Obtaining institutional regulatory approval—a process that runs in parallel with the design and development processes—is critical and can risk becoming the rate-limiting step. The evaluation of the product initially focuses on user experience, followed by clinical outcomes such as morbidity, mortality, or efficiency. With the majority of the platform built, the design and build times are dramatically reduced for clinic 2, and they continue to fall as the process becomes refined (eg, under 6 months for clinics 3-5) and occurs in parallel. MVP: minimum viable product.



HCD: Future Directions

Since the back-end infrastructure of an institution-wide platform is unified, only 1 set of corresponding regulatory approvals is needed, and this approach reduces the cost and time required to develop a front-end tool and allows multiple tools to be developed in parallel (Figure 3). However, given the number of medical specialties, clinical scenarios, disorders, and algorithms across a health care system, engaging in this intensive HCD process with each new clinic will not be cost-effective in the long term. Instead, a library of existing data sources and graphical interfaces could be generated, and clinicians (or patients, ie, in the event of a patient-facing version) could customize this library to design their own dashboards, thereby freeing programmers to concentrate on developing new modular interfaces and data sources. Generating more universal standards for describing clinical dashboards and their connections to APIs and EHRs could ease the deployment of dashboards across a wide range of health care platforms. Containerization, the Substitutable Medical Applications and Reusable Technologies (SMART) on Fast Healthcare Interoperability Resources (FHIR) API, and the Epic App Orchard (Epic Systems) represent important steps in this direction, but substantial scope for further standardization remains. The adoption of this type of adaptable clinical dashboard at scale would provide sufficient data for iteratively testing and improving performance, resulting in a second data-driven evaluation phase that focuses on surveys and click data. As the scale of data grows, especially across institutions, a third design phase that is based on both clinical outcomes and user experience will become possible.

Consideration 2: Technological Design

Common Approaches to the Technological Design of Digital Health Tools

The architecture of most digital health tools involves a connection among the back-end databases, middleware software algorithms that convert the data into useful knowledge, and front-end displays for users (Figure 2). Both single-clinic and centralized solutions are often hard-coded to represent a specific data source and visualization type, which slows the development of novel iterations and results in higher overall costs. A more efficient solution is to build a framework of reusable APIs that connects a multiplying number of data sources, computational algorithms, and modular visualization schematics and is adaptable and scalable to diverse types of medical data and clinical specialties.

Practical Considerations From the BRIDGE Experience

Overview of Practical Considerations

The BRIDGE platform was designed as a proof-of-principle MVP scaffold that could be developed efficiently and quickly but later refined and scaled up depending on its success and the collaborative opportunities generated. The HCD process made clear the following four key technical requirements: (1) it had

to permit access to a variety of data sources (ie, beyond the EHR), which could then be either displayed directly or processed through computationally intense algorithms [7]; (2) it needed to enable the visualization of these data in an intuitive and actionable manner, and this process needed to be embedded in the clinical workflow, so that it was not cumbersome for clinicians to access or operate; (3) following logically from the second requirement, it required the ability to launch directly from the EHR; and (4) it had to be as modular as possible to make iterative clinic-by-clinic customizations easier and more efficient to program.

Data Sources

Many data types contribute to precision care. To build a data foundation for BRIDGE that would best meet the needs of a variety of clinical use cases, we opted to include real-time clinical data from the EHR, minimally processed data from widely available data platforms (REDCap and Qualtrics [Qualtrics International Inc]) [16], data from institutional tools (eg, TabCAT [Tablet-Based Cognitive Assessment Tool; UCSF]) [17] and research databases [18], and complex data that either cannot be currently hosted in the standard EHR or require processing by complex analytics processing pipelines (Figure 2). For example, images from the Radiology PACS can be obtained ahead of time based on scheduled appointments, thereby allowing time for computationally intensive image processing pipelines to run prior to a patient appointment. Further innovations requiring advanced data processing include accessing expansive knowledge networks to compute precise clinical risk and treatment predictions [19]. As it would represent the convergence of so many sensitive data streams, BRIDGE required robust front-end and back-end architectures that were unified around security and hosted within the UCSF firewall.

Workflow Fit and EHR Integration

As a fundamental requirement for BRIDGE, to give clinicians actionable information during patient encounters, it had to launch directly from patients' records in the EHR (ie, Epic; Epic Systems) and pull their clinical data in real time (Figure 2). This resulted in a central technical decision to design BRIDGE as a SMART on FHIR application. Launching from the EHR resulted in additional clinical workflow benefits; discrete data could be collected at the point of care by using clinic-specific EHR Flowsheets and SmartForms (sharable across institutions), and data could then be pulled into clinical notes. Direct flowsheet data entry also allows BRIDGE to call and visualize discrete research data during clinic visits more efficiently. Enabling this launch functionality required interactions with the EHR development group and resources for funding their modifications to the EHR.

Modular Design

BRIDGE was designed to capitalize on a common language of clinical information flow through the creation of core widgets, or visualization modules, that can be adapted to an expanding array of clinical scenarios (Figure 2). At the time of BRIDGE MVP deployment, we had programmed the following four reusable core widgets: (1) longitudinal clinical course in the context of treatment, (2) cross-sectional metrics, (3) specialty-focused laboratory data, and (4) quantitative neuroimaging. Both the cross-sectional and longitudinal widgets allow patients' scores and metrics to be contextualized against a larger reference cohort that indicates both normal and abnormal values as well as percentile calculations, thus allowing a patient's clinical status to be interpreted by a clinician at a glance (Figure 2). We were able to convert existing precision medicine tools, such as the UCSF Multiple Sclerosis BioScreen longitudinal viewer [12] and the UCSF Brainsight magnetic resonance image processing and visualization tool [20] into these initial BRIDGE widgets. The configuration data for all viewers are stored by BRIDGE, which queries these data in real time and then renders the specified widgets and data sources for the clinician. Updates to the configuration can be made quickly when existing dashboards need to be adapted, thus enabling both ongoing user engagement and rapid deployment to meet the evolving needs of specific clinics. As we expand to other clinics, new widgets (eg, geolocation and genomics) that can be retroactively made available to existing clinics are being developed.

Technological Design: Future Directions

Two architectural changes can be made. The first is integration with a middleware platform. BRIDGE is currently connected to multiple data sources through direct API integrations, and connecting to additional APIs necessitates the modification of the codebase. Making use of a platform that aggregates APIs would reduce maintenance efforts and promote more stable platforms. Examples of such platforms already exist (eg, Human API [21]) and include EHR data. The second architectural change is creating a graphical user interface (GUI) that clinicians can use to create their own dashboards. Currently, dashboard configuration is done by the BRIDGE development team. Building a GUI that allows clinicians to configure and customize their dashboards would accelerate progress and allow clinicians without programming experience to access relevant data sources. Such an endeavor will likely require the integration of institution-wide platforms and centralized platforms, and such an integration will benefit both types of platforms. The resulting unified platforms would probably combine generic, cloud-based back-end and middleware components but be able to deliver the customized, clinic-specific, front-end dashboards designed by clinicians through the GUI. Overall, BRIDGE aims to augment—not supplant—the EHR; should an institution's visualization show clinical value, the institution could choose to maintain it in BRIDGE or integrate it into their EHR more permanently.

Innovations are also needed to improve data quality in the EHR, including tools that systematically flag likely data entry errors, simplify the correction of the EHR by a clinician, and ensure that corrections are distributed to all clinical tools. Finally, to demonstrate that these tools comply with the Health Insurance Portability and Accountability Act (HIPAA) or equivalent guidelines, a cross-institutional body that is responsible for testing and validating these solutions could be created. It might accelerate progress substantially by, for example, supporting cloud-based, HIPAA-compliant, off-the-shelf solutions to ease this data quality burden.

Consideration 3: Regulation and Policy

Launching a clinical application with real-time access to identified patient health data requires close institutional oversight and multiple stages of regulatory approval, especially in cases where clear institutional road maps or leadership structures are lacking due to the innovative nature of such applications.

Practical Considerations From the BRIDGE Experience

With regard to developing the BRIDGE MVP, the Epic EHR, and the SMART on FHIR application, technological capabilities were already available within our institution, but multiple security, privacy, technological, and compliance concerns had to be addressed. Specifically, authorizing an expandable, cross-specialty, modular platform rather than a domain- and clinic-specific tool was entirely novel. This necessitated parallel revisions to the approval process itself. Early in the design process, we set clear functional constraints that would reduce the barriers to institutional approval. Foremost among these were (1) conceptualizing BRIDGE as a clinical research tool that is custom designed with clinical specialists rather than as an institution-wide, enterprise-level clinical solution; (2) not requesting write access to the EHR (real-time read access was enabled); and (3) ensuring that data do not leave the institutional firewall. With an approved clinical research platform in place, the bar for institutional approval is substantially reduced for subsequent clinical dashboards that iterate on the initial design, reducing this multi-month process to a simple, clinic-specific sign-off (Figure 3). Further approval is required for applications that add novel functionalities or revisit one of the major systems constraints (eg, sending data to an external server).

Regulation and Policy: Future Directions

BRIDGE provides a mechanism for rapidly deploying and evaluating novel precision medicine algorithms and visualizations developed by clinical researchers [22-24] to evaluate their clinical benefit [25]. As the system expands and more clinical visualizations become the standard of care, medical centers may eventually choose to move the fundamental infrastructure of their institution-wide platforms from an MVP clinical research entity, such as BRIDGE, to a full, enterprise-level clinical system that delivers the same capabilities at a higher level of reliability [26,27]. This shift will be precipitated by a number of considerations, including the need for professional-level version control and releases; automated testing and quality control; the capacity for multilevel monitoring, logging, and auditing; and the ability to handle high user volumes without concurrency issues. The institution will also need to ensure that there is adequate personnel infrastructure behind the system to permit sustainable 24-hour user support and timely design and adaptation for new clinics. In the end, all stakeholders must be able to trust the reliability and clinical value of the final platform and the sustainability of the system supporting it [28]. For many such algorithms, moving along the continuum from clinical research to enterprise clinical care may well necessitate regulatory approval from the Food and Drug Administration Center for Devices and Radiological Health [29], as spelled out in their Digital Health Innovation Action Plan, and alignment with the international Software as a Medical Device guidelines through the International Medical Device Regulators Forum.

Consideration 4: Evaluation and Impact

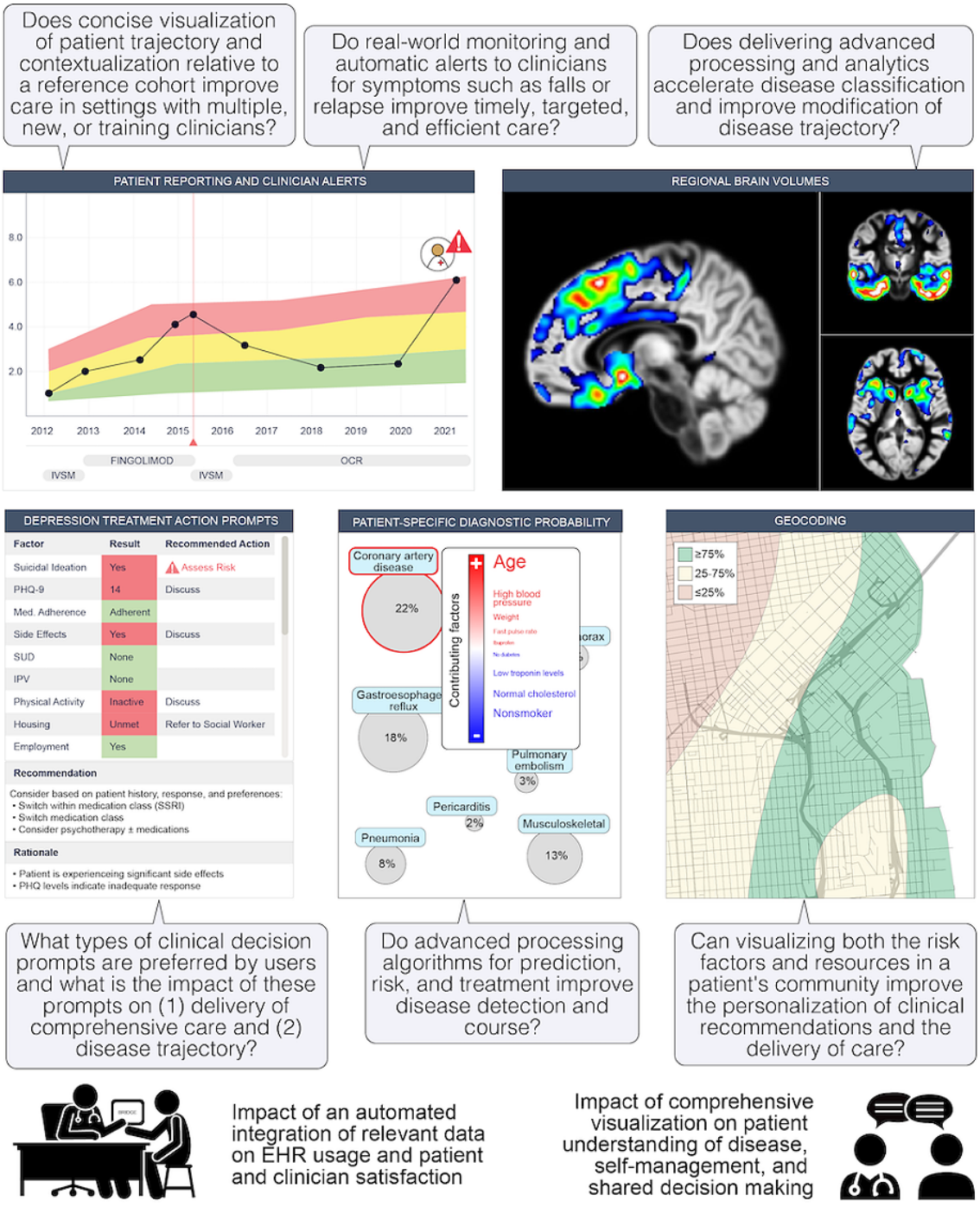
Pathway to Evaluation

Technological innovations in health care will ultimately be evaluated in terms of their impacts on patients, clinicians, data, and payors. In the near term, this requires the evaluation of a tool's interpretability and fidelity, that is, whether clinicians and patients like, understand, and use the tool and whether the use of the tool improves patients' experiences within the health system [15,28,30]. Making even the most complex algorithms visually digestible and actionable will be a key evaluation criterion [3]. To this end, for each BRIDGE dashboard, prior to measuring its clinical impact, we ensure that it meets key drivers of clinical adoption. We use the Health Information Technology Usability Evaluation Model [31] to evaluate at least 15 patients' and 8 clinicians' perceptions on the usefulness [32,33], ease of use [32,33], actionability [31], and likability [34] of each clinical dashboard. Low-scoring items (ie, <80% of respondents state "agree" for any given metric) engender another round of iterative development.

Evaluation and Impact: Future Directions

The impact of a dashboard like BRIDGE on clinical research and, eventually, care can be evaluated through in silico trials for answering a variety of clinical questions, as depicted in Figure 4. A near-term goal may be to compare users' preference for 2 types of symptom displays or to evaluate the impact of BRIDGE on workflow efficiency (eg, determining whether the use of the tool reduces the overall time spent on "clicking" through a patient's chart). Medium-term goals may be to refine a series of treatment action prompts that could yield a clinical decision support tool or to compare the effects of 2 different prediction algorithms on the risk of rehospitalization after a cardiac event. Long-term, altering clinical outcomes [2,25,27] that have obvious implications for health economics, such as reductions in the time to accurate diagnosis, rehospitalization, disability progression, morbidity, or death, will be directly relevant to an institution's assessment of a tool's utility.

Figure 4. Prototypes of clinical research enabled by BRIDGE. Clinical research applications include impacts on patient-doctor interactions and clinical workflows; the impact of monitoring patient-generated data, such as patient-reported outcomes and activity monitoring data; and the impact of delivering more advanced image processing and clinical algorithms (including prediction, prevention, and treatment algorithms) into the hands of clinicians. EHR: electronic health record; IPV: intimate partner violence; IVSM: intravenous solumedrol; MVP: minimum viable product; OCR: ocrelizumab; PHQ: Patient Health Questionnaire; PHQ-9: Patient Health Questionnaire-9; SSRI: selective serotonin reuptake inhibitor; SUD: substance use disorder.



Discussion

Determining whether big data analytics will truly disrupt clinical care depends on providing clinicians with access to the results of these analytics. In this paper, we describe one approach to overcoming the technical hurdle of making algorithms clinically available: the development of BRIDGE, an example of an institution-wide platform that allows for substantial clinic-specific customization. From the outset, BRIDGE was designed by intended users who worked closely with stakeholders, through an HCD process, to develop a structured and modular solution (Figure 2) that could be scaled and customized to specific clinic use cases in a cost- and time-efficient manner (Figure 3). The resulting platform addresses clinicians' requests to reduce data overload and more

precisely tailor the data that they use during clinical encounters. The lessons learned from building an institution-wide digital medicine platform include not only the importance of using HCD but also the importance of engaging with institutional partners and leadership early to collaboratively and transparently navigate through the long and arduous process of obtaining regulatory and security approval.

Based on our experiences, we propose that the development of similar platforms at other institutions is an efficient way to accelerate the testing of digital health algorithms in clinics. To reduce the burden of this undertaking, other academic clinical centers could use all or part of the BRIDGE platform code to create their own instances, especially if these centers used Epic, even though there would still be regulatory approval and software integration steps for making BRIDGE available within their EHRs. Additional developments could simplify this further, including sharing aspects of BRIDGE through centralized application stores, such as the Epic App Orchard, as well as creating centralized security audits and certifications that allow software to be vetted thoroughly once rather than vetting software at each new institution. Such centralization could be achieved by a federal initiative, a nationwide nonprofit society, or commercial vendors. For example, commercial vendors could provide institutions with centralized platforms that provide cloud-based computational resources, data access, security, and certification while clinicians and scientists develop dashboards and algorithms that run on these centralized platforms. BRIDGE provides a way to immediately develop and test these dashboards and algorithms in preparation for this future.

The potential of precision medicine will only be realized when the utility of the algorithms developed in this field can be evaluated at the point of care with real patients. Performing this testing requires substantial infrastructure development, which is hard to justify in the evaluation phase. Modular, scalable, institution-wide platforms, such as BRIDGE, represent one approach to resolving this catch-22 problem by providing an efficient mechanism for rapidly and cost-effectively deploying and evaluating new algorithms in clinics. Such a mechanism effectively serves as a bridge for translating research innovations into clinical tools.

Acknowledgments

The authors thank the funders (UCSF Weill Institute for Neurosciences), Dr Matthew State (for leadership of BRIDGE deployment in Psychiatry), Drs Ida Sim and Jason Satterfield (for iteration of a clinical action prompt tool), as well as Michael Schaffer (early BRIDGE architecture consultation) and Cosmo Mielke (w-map viewer in the UCSF Brainsight app).

Conflicts of Interest

RB receives research support from National Institutes of Health (NIH), California Initiative to Advance Precision Medicine, National Multiple Sclerosis Society (Harry Weaver Award), Hilton Foundation, and Sherak Foundation as well as Biogen, Novartis, and Roche Genentech. RB also receives scientific advisory board and consulting fees from Alexion, Biogen, EMD Serono, Genzyme Sanofi, Novartis, and Roche Genentech. BLM receives royalties from Guilford Press, Cambridge University Press, Johns Hopkins Press, and Oxford University Press and grant support from NIH and the Bluefield Project to Cure Frontotemporal Dementia. SLH serves on the scientific advisory boards of Accure, Alector, Annexon, and Molecular Stethoscope and the board of directors for Neurona. SLH has received travel reimbursement and writing support from Roche and Novartis for CD20-related meetings and presentations. KPR receives research funding from NIH, Quest Diagnostics, the Marcus Family Foundation, and the Rainwater Charitable Foundation. ES, PS, MG, SML, AB, and SJS have no conflicts of interest to declare.

References

1. Avram R, Olgin JE, Kuhar P, Hughes JW, Marcus GM, Pletcher MJ, et al. A digital biomarker of diabetes from smartphone-based vascular signals. Nat Med 2020 Oct;26(10):1576-1582 [FREE Full text] [doi: 10.1038/s41591-020-1010-5] [Medline: 32807931]
2. Cook DA, Enders F, Caraballo PJ, Nishimura RA, Lloyd FJ. An automated clinical alert system for newly-diagnosed atrial fibrillation. PLoS One 2015 Apr 07;10(4):e0122153. [doi: 10.1371/journal.pone.0122153] [Medline: 25849969]
3. Bean DM, Teo J, Wu H, Oliveira R, Patel R, Bendayan R, et al. Semantic computational analysis of anticoagulation use in atrial fibrillation from real world data. PLoS One 2019 Nov 25;14(11):e0225625. [doi: 10.1371/journal.pone.0225625] [Medline: 31765395]
4. Shah P, Kendall F, Khozin S, Goosen R, Hu J, Laramie J, et al. Artificial intelligence and machine learning in clinical development: a translational perspective. NPJ Digit Med 2019 Jul 26;2:69 [FREE Full text] [doi: 10.1038/s41746-019-0148-3] [Medline: 31372505]
5. Eaneff S, Obermeyer Z, Butte AJ. The case for algorithmic stewardship for artificial intelligence and machine learning technologies. JAMA 2020 Oct 13;324(14):1397-1398. [doi: 10.1001/jama.2020.9371] [Medline: 32926087]
6. Wiljer D, Hakim Z. Developing an artificial intelligence-enabled health care practice: Rewiring health care professions for better care. J Med Imaging Radiat Sci 2019 Dec;50(4 Suppl 2):S8-S14. [doi: 10.1016/j.jmir.2019.09.010] [Medline: 31791914]
7. Afzal M, Islam SMR, Hussain M, Lee S. Precision medicine informatics: Principles, prospects, and challenges. IEEE Access 2020 Jan 13;8:13593-13612 [FREE Full text] [doi: 10.1109/access.2020.2965955]

8. Grouin C, Deléger L, Rosier A, Temal L, Dameron O, Van Hille P, et al. Automatic computation of CHA2DS2-VASc score: information extraction from clinical texts for thromboembolism risk assessment. AMIA Annu Symp Proc 2011;2011:501-510 [FREE Full text] [Medline: 22195104]
9. Melgaard L, Gorst-Rasmussen A, Lane DA, Rasmussen LH, Larsen TB, Lip GYH. Assessment of the CHA2DS2-VASc Score in Predicting Ischemic Stroke, Thromboembolism, and Death in Patients With Heart Failure With and Without Atrial Fibrillation. JAMA 2015 Sep 08;314(10):1030-1038. [doi: 10.1001/jama.2015.10725] [Medline: 26318604]
10. CHA2DS2-VASc score for atrial fibrillation stroke risk. MDCalc. URL: https://www.mdcalc.com/cha2ds2-vasc-score-atrial-fibrillation-stroke-risk [accessed 2020-09-24]
11. Wang T, Oliver D, Msosa Y, Colling C, Spada G, Roguski Ł, et al. Implementation of a real-time psychosis risk detection and alerting system based on electronic health records using CogStack. J Vis Exp 2020 May 15(159):10.3791/60794 [FREE Full text] [doi: 10.3791/60794] [Medline: 32478737]
12. Gourraud P, Henry RG, Cree BAC, Crane JC, Lizee A, Olson MP, et al. Precision medicine in chronic disease management: The multiple sclerosis BioScreen. Ann Neurol 2014 Nov;76(5):633-642 [FREE Full text] [doi: 10.1002/ana.24282] [Medline: 25263997]
13. Design Kit: The human-centered design toolkit. IDEO. URL: https://www.ideo.com/post/design-kit [accessed 2022-01-06]
14. Matheson GO, Pacione C, Shultz RK, Klügl M. Leveraging human-centered design in chronic disease prevention. Am J Prev Med 2015 Apr;48(4):472-479. [doi: 10.1016/j.amepre.2014.10.014] [Medline: 25700655]
15. Bove R, Bruce CA, Lunders CK, Pearce JR, Liu J, Schleimer E, et al. Electronic health record technology designed for the clinical encounter: MS NeuroShare. Neurol Clin Pract 2021 Aug;11(4):318-326. [doi: 10.1212/CPJ.0000000000000986] [Medline: 34484932]
16. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
17. Possin KL, Moskowitz T, Erlhoff SJ, Rogers KM, Johnson ET, Steele NZR, et al. The brain health assessment for detecting and diagnosing neurocognitive disorders. J Am Geriatr Soc 2018 Jan;66(1):150-156 [FREE Full text] [doi: 10.1111/jgs.15208] [Medline: 29355911]
18. University of California, San Francisco MS-EPIC Team:, Cree BAC, Gourraud P, Oksenberg JR, Bevan C, Crabtree-Hartman E, et al. Long-term evolution of multiple sclerosis disability in the treatment era. Ann Neurol 2016 Oct;80(4):499-510 [FREE Full text] [doi: 10.1002/ana.24747] [Medline: 27464262]
19. Nelson CA, Bove R, Butte AJ, Baranzini SE. Embedding electronic health records onto a knowledge network recognizes prodromal features of multiple sclerosis and predicts diagnosis. J Am Med Inform Assoc 2021 Dec 16:ocab270. [doi: 10.1093/jamia/ocab270] [Medline: 34915552]
20. Memory and Aging Center. University of California, San Francisco. URL: https://memory.ucsf.edu/ [accessed 2022-01-24]
21. Human API. Human API. URL: https://www.humanapi.co/ [accessed 2022-01-24]
22. Olgin JE, Lee BK, Vittinghoff E, Morin DP, Zweibel S, Rashba E, et al. Impact of wearable cardioverter-defibrillator compliance on outcomes in the VEST trial: As-treated and per-protocol analyses. J Cardiovasc Electrophysiol 2020 May;31(5):1009-1018. [doi: 10.1111/jce.14404] [Medline: 32083365]
23. Norgeot B, Glicksberg BS, Trupin L, Lituiev D, Gianfrancesco M, Oskotsky B, et al. Assessment of a deep learning model based on electronic health record data to forecast clinical outcomes in patients with rheumatoid arthritis. JAMA Netw Open 2019 Mar 01;2(3):e190606 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.0606] [Medline: 30874779]
24. Hong JC, Eclov NCW, Dalal NH, Thomas SM, Stephens SJ, Malicki M, et al. System for High-Intensity Evaluation During Radiation Therapy (SHIELD-RT): A prospective randomized study of machine learning-directed clinical evaluations during radiation and chemoradiation. J Clin Oncol 2020 Nov 01;38(31):3652-3661. [doi: 10.1200/JCO.20.01688] [Medline: 32886536]
25. Panch T, Pollard TJ, Mattie H, Lindemer E, Keane PA, Celi LA. "Yes, but will it work for patients?" Driving clinically relevant research with benchmark datasets. NPJ Digit Med 2020 Jun 19;3:87 [FREE Full text] [doi: 10.1038/s41746-020-0295-6] [Medline: 32577534]
26. Parikh RB, Obermeyer Z, Navathe AS. Regulation of predictive analytics in medicine. Science 2019 Feb 22;363(6429):810-812 [FREE Full text] [doi: 10.1126/science.aaw0029] [Medline: 30792287]
27. Choudhury A, Asan O. Role of artificial intelligence in patient safety outcomes: Systematic literature review. JMIR Med Inform 2020 Jul 24;8(7):e18599 [FREE Full text] [doi: 10.2196/18599] [Medline: 32706688]
28. Asan O, Bayrak AE, Choudhury A. Artificial intelligence and human trust in healthcare: Focus on clinicians. J Med Internet Res 2020 Jun 19;22(6):e15154 [FREE Full text] [doi: 10.2196/15154] [Medline: 32558657]
29. Allen B. The role of the FDA in ensuring the safety and efficacy of artificial intelligence software and devices. J Am Coll Radiol 2019 Feb;16(2):208-210. [doi: 10.1016/j.jacr.2018.09.007] [Medline: 30389329]
30. Shinners L, Aggar C, Grace S, Smith S. Exploring healthcare professionals' understanding and experiences of artificial intelligence technology use in the delivery of healthcare: An integrative review. Health Informatics J 2020 Jun;26(2):1225-1236 [FREE Full text] [doi: 10.1177/1460458219874641] [Medline: 31566454]

31. Brown 3rd W, Yen PY, Rojas M, Schnall R. Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology. J Biomed Inform 2013 Dec;46(6):1080-1087 [FREE Full text] [doi: 10.1016/j.jbi.2013.08.001] [Medline: 23973872]
32. Yen P, Sousa KH, Bakken S. Examining construct and predictive validity of the Health-IT Usability Evaluation Scale: confirmatory factor analysis and structural equation modeling results. J Am Med Inform Assoc 2014 Oct;21(e2):e241-e248 [FREE Full text] [doi: 10.1136/amiajnl-2013-001811] [Medline: 24567081]
33. Schnall R, Cho H, Liu J. Health Information Technology Usability Evaluation Scale (Health-ITUES) for usability assessment of mobile health technology: Validation study. JMIR Mhealth Uhealth 2018 Jan 05;6(1):e4 [FREE Full text] [doi: 10.2196/mhealth.8851] [Medline: 29305343]
34. Mathews SC, McShea MJ, Hanley CL, Ravitz A, Labrique AB, Cohen AB. Digital health: a path to validation. NPJ Digit Med 2019 May 13;2:38 [FREE Full text] [doi: 10.1038/s41746-019-0111-3] [Medline: 31304384]

Abbreviations

API: application programming interface
EHR: electronic health record
FHIR: Fast Healthcare Interoperability Resources
GUI: graphical user interface
HCD: human-centered design
HIPAA: Health Insurance Portability and Accountability Act
MVP: minimum viable product
NIH: National Institutes of Health
PACS: Picture Archiving and Communication System
REDCap: Research Electronic Data Capture
SMART: Substitutable Medical Applications and Reusable Technologies
TabCAT: Tablet-Based Cognitive Assessment Tool
UCSF: University of California, San Francisco

Edited by G Eysenbach; submitted 29.10.21; peer-reviewed by M Afzal; comments to author 25.11.21; revised version received 17.12.21; accepted 22.12.21; published 15.02.22.

Please cite as:
Bove R, Schleimer E, Sukhanov P, Gilson M, Law SM, Barnecut A, Miller BL, Hauser SL, Sanders SJ, Rankin KP
Building a Precision Medicine Delivery Platform for Clinics: The University of California, San Francisco, BRIDGE Experience
J Med Internet Res 2022;24(2):e34560
URL: https://www.jmir.org/2022/2/e34560
doi:10.2196/34560
PMID:35166689

Viewpoint

Lessons Learned From Beta-Testing a Facebook Group Prototype to Promote Treatment Use in the "Connecting Alaska Native People to Quit Smoking" (CAN Quit) Study

Pamela S Sinicrope[1], MPH, DPH; Colleen D Young[2], BA; Ken Resnicow[3], PhD; Zoe T Merritt[4], MBA; Clara R McConnell[5], BA; Christine A Hughes[1], BS; Kathryn R Koller[4], PhD; Martha J Bock[1], BS; Paul A Decker[1], MS; Christie A Flanagan[4], MPH; Crystal D Meade[5], BS; Timothy K Thomas[4], MD; Judith J Prochaska[6], MPH, PhD; Christi A Patten[1], PhD

[1]Behavioral Health Research Program, Department of Psychiatry and Psychology, Mayo Clinic, Rochester, MN, United States
[2]Division of Consumer Communications, Social and Digital Innovation, Mayo Clinic Connect, Rochester, MN, United States
[3]Department of Health Behavior and Health Education, School of Public Health, University of Michigan, Ann Arbor, MI, United States
[4]Clinical & Research Services, Division of Community Health Services, Alaska Native Tribal Health Consortium, Anchorage, AK, United States
[5]Wellness and Prevention, Division of Community Health Services, Alaska Native Tribal Health Consortium, Anchorage, AK, United States
[6]Stanford Prevention Research Center, Department of Medicine, Stanford University, Stanford, CA, United States

Corresponding Author:
Pamela S Sinicrope, MPH, DPH
Behavioral Health Research Program
Department of Psychiatry and Psychology
Mayo Clinic
200 1st St SW
Rochester, MN, 55905
United States
Phone: 1 507 284 2511
Email: Sinicrope.Pamela@mayo.edu

Abstract

Social media provides an effective tool to reach, engage, and connect smokers in cessation efforts. Our team developed a Facebook group, CAN Quit (Connecting Alaska Native People to Quit smoking), to promote use of evidence-based smoking cessation resources for Alaska Native people living in Alaska, which are underused despite their effectiveness. Often separated by geography and climate, Alaska Native people prefer group-based approaches for tobacco cessation that support their culture and values. Such preferences make Alaska Native people candidates for social media–based interventions that promote connection. This viewpoint discusses the steps involved and lessons learned in building and beta-testing our Facebook group prototype, which will then be evaluated in a pilot randomized controlled trial. We describe the process of training moderators to facilitate group engagement and foster community, and we describe how we developed and tested our intervention prototype and Facebook group. All parts of the prototype were designed to facilitate use of evidence-based cessation treatments. We include recommendations for best practices with the hope that lessons learned from the CAN Quit prototype could provide a model for others to create similar platforms that benefit Alaska Native and American Indian people in the context of smoking cessation.

(J Med Internet Res 2022;24(2):e28704) doi:10.2196/28704

KEYWORDS

Web 2.0; social media; Facebook; Alaska Native; American Indian; Alaska; smoking; cessation; cancer prevention; Quitline; mobile phone

Introduction

Background

Social media is a powerful tool for reaching, engaging, and connecting smokers in cessation efforts. This could be especially true for Alaska Native people, who have the highest smoking prevalence in any US racial or ethnic subgroup [1]. Furthermore, Alaska Native residents of Alaska have a smoking prevalence more than double than that of Alaskan White residents (42% vs 17%), [2,3] but are difficult to reach with traditional face-to-face interventions because of their remote geography [4]. More than half of Alaska Native people live in more than 200 small communities (populations averaging 500-1000 total residents). Most of these communities are located off the road system and are accessible year-round only by a small plane, making them both geographically and socially remote.

Research indicates that internet (web)-based interventions are effective for smoking cessation and overcoming barriers related to travel; however, they are also associated with low use [5,6]. Conversely, social media has been shown to promote access and engagement among underserved, diverse populations with evidence-based content and peer support [7-9]. In contrast to individual-based treatments, social media platforms could lead to greater adoption and sustainability by encouraging collaborative efforts across adult generations that resonate with the Alaska Native cultural value of interdependence, defined as relationship oriented and collaborative in decision-making and lifestyle changes [10-12]. It is also important to note that traditional Alaska Native lifestyles, cultures, and holistic worldviews are intertwined and extend beyond mere social collectivism to include the interconnectedness of people with nature, the environment, and all elements of the universe [13,14]. Furthermore, Facebook is the dominant social networking platform used by 68% of US adults, more than double the proportion on Twitter (21%), Instagram (28%), Pinterest (26%), and LinkedIn (25%). Of Facebook users, 75% engage with the site daily. Facebook use indicates similarly high rates of engagement by sex (75% of females and 83% of males use Facebook), income (84% engagement among those earning <US $30,000/year vs 77% for those reporting >US $74,000/year), and age groups (88%, 18-29 years; 84%, 30-49 years; and 72%, 50-74 years), except for slightly less use (62%) among those aged ≥65 years. Of Alaskans, 91% reported having mobile broadband internet access [15-19], and surveys showed that the Alaska Native people in both rural and urban areas of Alaska use their smartphones to access social media, specifically Facebook [17,18,20].

On the basis of the above reasons, Alaska Native people, who value connection and community and represent an underserved health disparities group with significant barriers to receiving cessation treatment [21], may be optimal candidates for a social media platform where they can learn about evidence-based cessation resources in a group setting with other Alaska Native people also interested in quitting smoking. Furthermore, a Facebook intervention represents a potential platform for study recruitment, as Facebook is a primary source for social networking [17,18].

Objective

The overall purpose of the CAN Quit (Connecting Alaska Native People to Quit Smoking) intervention study is to promote use of evidence-based smoking cessation resources available to Alaska Native people state-wide using social media. In the first 2 phases, we evaluated content for inclusion in the study using qualitative and quantitative methods [22,23]. In this third phase, our goal is to develop and beta-test the intervention prototype. The purpose of this paper is to highlight the results of the beta-testing phase of this 4-phase study and share lessons learned that could be applied by others interested in similar health behavior interventions [22].

What Is the CAN Quit Study?

CAN Quit [22] is an ongoing 4-phase study to iteratively develop and pilot-test a culturally tailored social media–delivered intervention (via a secret Facebook group) to promote evidence-based smoking treatment uptake and cessation among Alaska Native people who smoke. The need for CAN Quit arose from a long-term collaboration between the Alaska Native Tribal Health Consortium (ANTHC) and Mayo Clinic. The goal of CAN Quit was to use peer support via social media to develop a scalable and sustainable intervention that could enhance use and reach of existing and effective cessation services offered to Alaska Native people through ANTHC, other Tribal cessation resources, and resources available through the Alaska State Quitline. If effective, we anticipate that ANTHC could manage and sustain CAN Quit as part of its tobacco cessation program and be moderated by its tobacco cessation counselors.

CAN Quit uses a culturally based digital storytelling approach [24-27], with content adapted from the effective Centers for Disease Control and Prevention (CDC) Tips from Former Smokers campaign [27] and from the ANTHC library of digital stories. Digital stories previously created by Alaska Native people contained their voices, pictures, and stories about tobacco cessation. We incorporated cultural variance and surface or deep structure frameworks [28] to address the influence of culture in designing health messages. The Mayo Clinic and Alaska Area institutional review boards and ANTHC Board of Directors' Health Research Review Committee reviewed and approved all study phases.

In the first 2 study phases (described in the studies by Sinicrope at al [22] and Merculieff et al [23]), we gathered feedback first qualitatively and then quantitatively from Alaska Native people who smoke recruited state-wide primarily through targeted paid Facebook ads. We asked about the suitability of content, including digital stories and photos with text provided by CDC Tips and ANTHC. Similar to Tips, taglines to all content included a call to action for tobacco treatment by providing the (1) State of Alaska toll-free Quitline number, (2) URL to regional Tribal tobacco cessation program websites, and (3) URL to the smokefree.gov quit smoking resources website. In phases 1 and 2, through analysis of qualitative interviews with Alaska Native people who smoke and stakeholders (Alaska Native tobacco counselors) who viewed existing text, image, and video content, we found that participants preferred content that included Alaska Native people engaged in Alaska Native

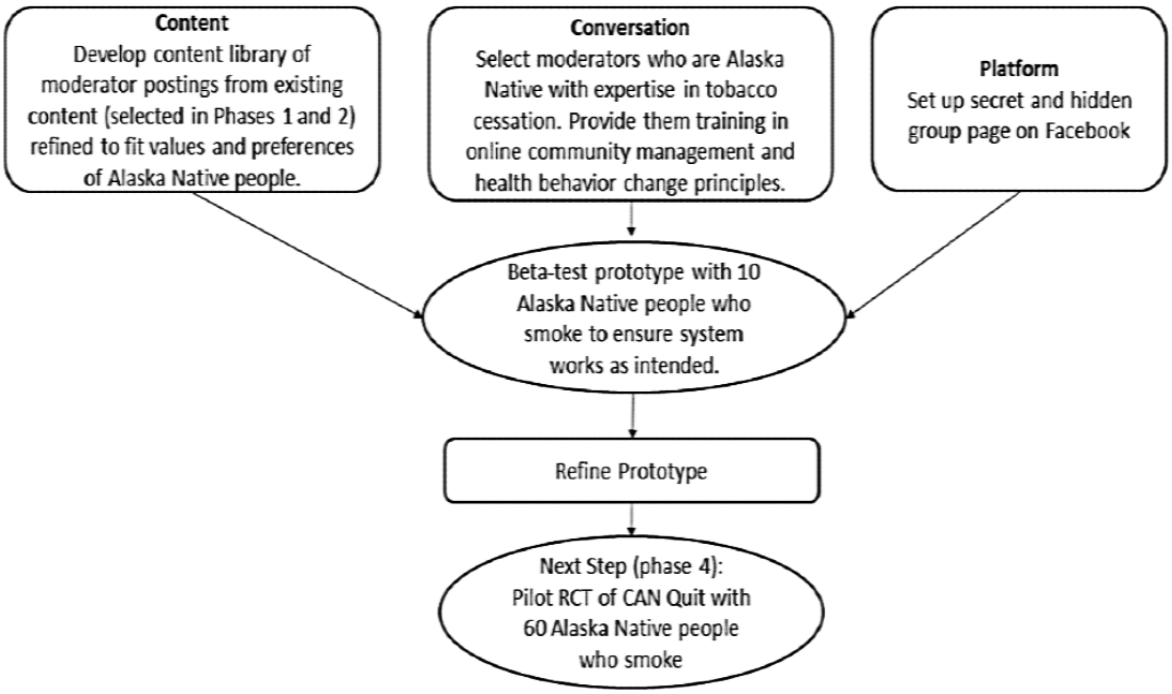
activities and were motivated to change smoking behavior when presented with images reflecting Alaska Native values of family, children, and community. Participants also preferred direct and honest storytelling styles told by other Alaska Native people without "sugar-coating" [23].

Prototype Components

Overview

Through this formative work, we developed an intervention prototype (phase 3) that included the following: (1) a Facebook group interface populated with culturally appropriate images and text outlining the purpose of the group, (2) a content library of images with text and digital stories (video) selected with sample moderator postings, and (3) a training plan for 2 Alaska Native study coordinators to moderate the Facebook group. Guidelines from Pagoto et al [29] informed the development of the components of our prototype and the training plan. Figure 1 provides an overview of the intervention prototype.

Figure 1. Overview of the CAN Quit (Connecting Alaska Native People to Quit Smoking) intervention prototype. CAN Quit is a 4-phase study. This figure highlights how phase 3 fits into the overall study design. RCT: randomized controlled trial.



Facebook Group Interface

We created community guidelines specific to CAN Quit to establish and maintain a supportive, inclusive, and respectful community. These were built on standard group conduct guidelines provided by Facebook when creating a group on the platform. The study team reviewed the standard group conduct guidelines together and elaborated on them to be specific to the CAN Quit group. For example, the group elements were as follows: (1) a name and cover photo for the group; (2) a description of the group's purpose, including contact information and a link to evidence-based quitting resources; and (3) an *about* section describing the group, its purpose, and five basic guidelines: be respectful and courteous, no hate speech or bullying, protect yourself and the group's privacy, no commercial posts or promotions and spam, and be careful about giving out medical advice. Privacy settings were established as *closed* (ie, only members could see who was in the group and what they post) and *secret* (ie, only members could find this group). Similar to CDC Tips content, prominent links to evidence-based quit smoking resources were located on all content and on the cover page.

Content Library

The team developed a content library organized similar to a conventional treatment manual, available only to the moderators; however, once posted, Facebook stored the content on the page history and in the photo library. We organized pivotal discussion points by content, taglines, and sample text so that moderators could initiate discussions to promote quitting. The final 64-page library (digital and written) included the following: (1) welcome posts, group description, guidelines, and contact information; (2) options for Facebook cover photos; (3) 64 posts (images and text or video that included a tag line to call the Quitline or use Tribal quitting resources, similar to the CDC Tips campaign); (4) guidelines for handling inappropriate or misinformation (eg, deleting unrelated or false posts and removing or blocking participants breaking group guidelines); and (5) a list of quitting resources, including the State of Alaska Quitline and regional Tribal resources.

We organized posts by subject (ie, quitting reasons and sources of quitting support), and all content included sample text the moderators can modify to fit with or be responsive to current group conversations. Figure 2 includes 2 images, one is the original CDC post that features Rebecca, a 57-year-old White

woman sitting on the sofa in her living room with the tag line "Quitting isn't about what you give up. It's about what you get back." The second adapted image includes the same tag line, but the image is that of an Alaska Native grandfather holding a baby, which resonates with the Alaska Native value of family and its importance as a motivator for smoking cessation. Sample text in the form of the question, "What are you hoping to gain from becoming smoke-free?" accompanied the post in which the moderator could add in to generate a discussion.

Figure 2. Example of Centers for Disease Control and Prevention Tips post and how it was refined for CAN Quit (Connecting Alaska Native People to Quit Smoking).

Moderators

The success of an online community is heavily dependent on the skills of moderators. Embedded in this skill is their familiarity with the topic (tobacco cessation), library content, understanding of and appreciation for Alaska Native cultural values, and ability to facilitate communication in a Facebook group. Recognizing this, we stressed hands-on scenario-based training and working directly with participants in each phase of the study so that moderators could continually develop their skills and become comfortable with working in the online community. Both Facebook moderators are trained tobacco specialists (TTSs) and Alaska Native employees of the ANTHC Tobacco Cessation Program in Anchorage, Alaska. Each received TTS training in motivational interviewing (MI) and additional training (as described in the section *Training Plan* below) to moderate CAN Quit. Moderators posted content every 2 to 3 days to prompt and reinforce participant posts and discussions (eg, sharing stories of tobacco use and quitting). Although content was prepared in advance in anticipation of discussions that would emerge, moderators also posted when relevant to guide and direct conversation and to provide timely content in response to member directions and needs.

Training Plan

Overview

A platform such as Facebook provides space for creating groups, but building a successful online community where members actively participate, engage with one another, and develop relationships requires an enabler (ie, a moderator) with organized strategic community management skills [30-33]. In addition, creating an online [30] community that serves Alaska Native people also requires that community managers and content developers pay attention to cultural norms and ways of interacting [31]. Therefore, the research team prioritized a moderator training using a scenario-based approach [34], providing moderators with skills to promote engagement, communication, and health behavior change (ie, use of evidenced-based quitting resources), whereas moderators provided feedback on how scenarios could be consistent with Alaska Native culture. The overall goal was to optimize and tailor the training specific to the needs of the CAN Quit study using *hands-on* instruction with real-life situations or scenarios relevant to learners [35]. In this instance, we created a CAN Quit practice Facebook group with posts and mock situations for moderators to practice crafting responses.

Two phases of training for moderators covered essential elements of online community management and moderator practices in part 1 and MI, tailoring, and cultural adaptation in part 2. The moderators actively participated in developing the training, providing their expertise and background as Alaska Native TTSs working with Alaska Native people.

Online Community Management and Moderator Practices Training

Over the course of 12 training hours, moderators acquired skills in strategic community management and applicable moderator practices required to build and maintain a thriving online health community (presented by CY). The training included the

following: (1) a pretraining questionnaire; (2) a 90-minute web-based introduction to online community management (see the section *Online Training*); (3) a 2-day in-person workshop (10.5 hours) with hands-on practice moderating and time for feedback and cultural tailoring; and (4) ongoing coaching through observing moderation practices in situ.

Pretraining Questionnaire

The pretraining questionnaire asked moderators four questions: (1) What do you hope the CAN Quit Facebook group will achieve? (2) As a moderator, what do you anticipate will be your single biggest challenge? (3) As a moderator, what type of interaction in the community would make you feel rewarded? and (4) What does success look like? Answers to these questions helped tailor the initial web-based training session.

Online Training

This 90-minute training was divided into 2 sessions. Session 1 introduced the fundamentals of community management, such as defining online community and understanding why people join patient support networks and what motivates them to return. The session addressed the interplay among the three ingredients for success: growth, activity, and sense of community. It also explored strategies and tactics supporting these items throughout the life cycle of an online community. The session examined strategic approaches for program inception: (1) welcoming new members and integrating them into the community; (2) ensuring every post receives a response; (3) connecting members; (4) keeping people talking; (5) asking questions that encourage self-disclosure and build trust; and (6) discovering members' knowledge, capacity, and strengths.

During the subsequent session, trainees posed questions to the community management expert or trainer in *Ask Me Anything* style. In this session, questions frequently include the following: (1) How do you engage inactive members? (2) How do you encourage meaningful interactions? (3) What do you do about members who do not follow community guidelines? Although these questions were addressed in brief during the *Ask Me Anything*, we also addressed them in the in-person workshop.

In-Person Workshop

A strategic community management plan ensures that the achievements of the online community and its members align with the overarching goal of sponsoring institutions. In this 2-day workshop, moderators and trainers developed a strategic community management plan (Figure 3), followed by hands-on practice applying strategies consistent with situational learning strategies [34]. The trainer (CY) relies on the best practices by Millington [30] adapted to a health setting in this study and in similar work with the online community Mayo Clinic Connect [32,36,37].

Figure 3. CAN Quit (Connecting Alaska Native People to Quit Smoking) strategic community plan.

Goal	Decrease smoking rates among Alaska Native people				
Objectives	**Convert** first recruits from passive to active participating members.			**Connect** members and get them to respond to other members, encourage them, and share their experiences.	
Strategies	Make members feel comfortable sharing their situation openly and asking questions	Make members feel proud as they set goals, show progress and achieve goals.	Make members feel accepted to come back to the community if they have a set back.	Make members feel joy in helping and being there for other members.	Make members feel confident they have useful experience and expertise to share. Make them feel comfortable asking questions of others.
Tactics	Ask questions to get more info about the member and to discover their knowledge.	Get members to share a target or goal. Celebrate milestones.	Find the positive in the setback to set a new goal.	Foster and reinforce positive connecting behavior.	Tag members and tell them why you did.

Strategic Planning Framework

The first day of the workshop focused on how to establish a framework for building an online community that includes the building blocks (goal, objectives, strategy, and tactics). The goal was defined as the direct value the organization gets from the community. Objectives are what members need to do to get this value. Strategy involves the emotion moderators amplify to get members to perform this behavior. Tactics are the things moderators do to amplify this emotion [30].

Importantly, the workshop emphasized that each action moderators take in the online community should be meaningful: (1) give new members confidence to make a first post, get help, and return to help others; (2) demonstrate the value of the CAN Quit group, a community where no question goes unanswered and people are supportive; and (3) amplify a sense of personal progress, success, and growing confidence in each member as they reach milestones and goals, or return for encouragement and support if a setback is experienced. Trainees were asked to reflect on the following questions: (1) What do you want members of the CAN Quit group to do? and (2) What do you have to do to make that happen? In the inception stage of this limited enrollment group, tactics were focused on moderator actions promoting activity and a sense of community objectives.

The remainder of the 2-day workshop concentrated on specific moderator approaches to support the above tactics. Although the moderator approaches remained platform agnostic, specific Facebook tools and their optimization were also discussed and a hidden *Practice CAN Quit* group was created to practice skills. Trainees learned and practiced the following: (1) *Setting up the group*: What do you want members to see when they visit the group? What 1 action do you want them to take? (2) *Getting*

the conversations started: How do you create conversations members want to participate in? (3) *Managing the tough stuff:* How do you deal with members who do not follow the community guidelines? (4) *Effectively using private messaging:* When should you use and not use private messaging? and (5) *Managing the group when it becomes active:* When should a moderator stay out of the way? How do you transition from micro to macro interactions? The team spent additional time discussing ways to monitor or track posts, strategy, and tactics to determine which may need adjustment.

The training wrapped up with a reminder of the importance of self-care. From the experience of the trainer (CY), reading all posts, responding with empathy, and helping members succeed can be emotionally taxing. The group addressed good self-care practices, including recognizing when to unplug from the group and taking time off and reaching out to fellow moderators and the research team for difficult situations or decisions.

Ongoing Coaching Through Observing Moderation Practices In Situ

Following the training, trainer CY provided moderators feedback biweekly, as they put the newly acquired skills into practice. They were given tips on how posts and interactions could be tweaked and recognizing opportunities to take action to promote further engagement. Moderators also reached out to other members of the research team with questions and concerns.

Effective Communication Strategies and Cultural Adaptation Training

In this training, moderators received approximately 10 hours of web-based training, divided across 2 days, approximately 5 months apart (July and December 2019). The trainer (KR) addressed the following: (1) theories of behavior change and tailored communications, (2) application of MI techniques for social media interventions, and (3) cultural adaptation of smoking cessation messages.

Theory of Behavior Change and Tailored Communications

The group discussed several behavior change conceptual models, including self-determination theory [38] and cognitive behavioral therapy [39,40]. The author KR also presented principles of cultural and psychosocial tailoring of health communications [28,41], at both the individual and group levels. He then applied theoretical models to cases directly addressing smoking cessation behavior in an Alaska Native population.

MI Techniques in a Social Media Intervention

MI is a client-centered counseling approach that has been tested in hundreds of studies, including numerous studies on smoking cessation [42]. This aspect of the training emphasized how to use key MI principles (eg, autonomy support) and strategies (rolling with resistance and eliciting change talk) of MI in the context of this social media intervention. The group discussed and practiced strategies for writing effective questions and reflective responses and for acknowledging and eliciting change talk, such as the following:

- "Sounds like you are moving toward giving up smoking."
- "Your post indicates you have lots of good reasons to quit, such as your kids, your community, and your health."
- "You are proud of yourself for staying quit for more than two weeks."
- "Although you are not quite ready to quit yet, it sounds like you are moving in that direction. You are getting close."

Cultural Adaptation of Smoking Cessation Messages

Moderators received training on incorporating the Alaska Native cultural content into their Facebook posts and responses. This included a review of Alaska Native core values and the practice of generating messages that linked smoking cessation to these values. For example, the team brainstormed ways that respect for nature, community, humility, and mindfulness could be linked to cessation messages. The team also explored ways to post polls on the Facebook page that queried users' core values, which they could then use to generate tailored messages about how their values might drive their efforts to reduce their tobacco use.

Training applied web-based exercises and actual or simulated Facebook posts in a scenario-based learning approach to enable moderators to master concepts and build messaging skills [34]. When given posts, the moderators generated responses in real time using the principles and strategies presented. For example, moderators received the following participant post:

> *I know quitting is going to be hard but I want to show my kids that I'm strong and I want to be around for them when they're older and have their own kids. I want to be there for my family.*

For this case, the response included the strategy of rolling with resistance around the perceived difficulty of quitting and use of a reflective statement capturing the meaning of quitting for this participant; that is, change talk about her desire to be there for her family, such as using a reflective state, for example, "on one hand you are scared that quitting will be challenging," as well as reflecting the meaning of quitting back to the participant that is her change talk, for example, "but quitting would send them a message that you care about them and are willing to do what's needed to be there for them."

Beta Test

Overview, Sample, and Methods

Upon completion of all training and 2 weeks of practice, we beta-tested the intervention prototype with our 2 moderators and a sample of 10 Alaska Native adults living in Alaska who smoke cigarettes. As in our prior study phases described previously [22], Alaska Native people who smoke were recruited state-wide primarily via targeted paid Facebook ads that included an image and short text consistent with Facebook's advertising guidelines based on the following criteria: aged >19 years, self-reported Alaska Native race or ethnicity, and keywords related to tobacco use. To be eligible to participate, participants had to be living in Alaska, aged ≥19 years, have smoked >1 cigarette per day over the past 7 days with cigarettes as the main tobacco product used, and be considering or willing to make a quit attempt. They also had to have access to broadband (high-speed) internet on a mobile phone, at home,

work, or other locations, and a Facebook account or be willing to set one up. Finally, they could not have been enrolled in a cessation program or using pharmacotherapy to stop smoking during the last 3 months. Prior work suggested 10 people as the minimum for optimal engagement in social media [43,44].

The purpose of this phase was to provide participants 30 days of exposure to the CAN Quit Facebook group and obtain their feedback to (1) ensure the system worked as intended, (2) identify technical issues, and (3) facilitate program refinements in preparation for the pilot randomized controlled trial (RCT). We administered a web-based survey at baseline to collect demographic, smoking behavior, and readiness to quit information. At completion of the 1-month beta test, we assessed readiness to quit, quit attempts, and use of cessation resources. In addition, we also assessed the acceptability of the CAN Quit group platform via the social media usability scale, which measures perceived ease of use and usefulness (3 items each), ease of learning (1 item), and satisfaction (6 items), rated on a 5-point scale (1—strongly disagree, 5—strongly agree) [45]. We also asked open-ended questions to provide feedback on modifications to the prototype. Participants received a US $25 gift card for their time after each assessment.

Analyses

We summarized the sample characteristics and usability measures using descriptive statistics (means, SDs, percentages, and frequencies). We analyzed the responses to open-ended questions using content analysis [46].

Results

Baseline

Of the 10 Alaska Native adults (mean age 35, SD 8.18; range 23-52 years; 9/10, 90% female; 7/10, 70% living in a rural area; 6/10, 60% employed for pay; and all had a high school degree or higher), the average number of cigarettes smoked per day was 10.2 (SD 7.5; range 2-25) at baseline. Furthermore, 60% (6/10) reported a high level of readiness to quit cigarettes (mean 7.2, SD 2.7; range 2-10), and 50% (5/10) reported readiness to use smoking cessation treatment (mean 7.1, SD 2.6; range 0-10) (data not shown).

One-Month Follow-Up

Among the 10 participants, 6 reported at least one quit attempt and 2 quit smoking (both called the Quitline). Feedback on the Facebook group revealed an overall social media usability score of 4.1 (SD 0.58; range 3.1-5.0) for the 13 items reported. Mean categorical scores were as follows: for usefulness (3 items), 3.7 (SD 0.66; range 3.0-5.0); for ease of use (3 items), 4.5 (SD 0.61; range 3.0-5.0); for ease of learning (1 item), 4.7 (SD 0.67; range 3.0-5.0); and for satisfaction (6 items), 4.0 (SD 0.71; range 3.1-5.0).

Facebook Page Activity Summary

We conducted a beta test in the fall of 2019. Moderators posted 19 unique posts from the content library, 1 participant created a post with a photo of their children as a reason for quitting, and the cover page was posted. Participants posted a total of 35 reactions (eg, all *likes* and *loves*), and 130 comments were generated across the posts, with 57.7% (75/130) of comments coming from moderators and 42.3% (55/130) from users. Of the 10 pilot testers, 9 posted 54 comments in all. The number of comments per person ranged from 1 to 24 (median 5). For the 2 people who reported quitting smoking, 1 posted 15 comments and the other posted 2. There was only 1 instance where a participant shared content outside the group on a personal Facebook page. When this occurred, the moderator sent the participant a private message, redirecting the participant to the group rules. There were no other boundary violation incidents (data not shown).

Participant Feedback on Group

Participants and moderators were asked to provide feedback on how to improve moderator posts as well as the Facebook group overall (Table 1). Four areas of response emerged: (1) providing incentives for participation, (2) clarifying the group purpose, (3) making posts more interactive and specific to rural Alaska Native people, and (4) encouraging group members to post more frequently. Moderators expressed that they felt prepared to use the skills from their training. They expressed challenges associated with working with the Facebook platform to ensure they did not miss comments because Facebook does not organize posts in chronological order in groups. They also expressed challenges associated with keeping track of multiple users across posts and found that maintaining a spreadsheet of group members and their stories helped to ensure that they responded appropriately to posts. One of the biggest challenges was responding to participants who shared traumatic or personal information. However, they expressed that training in MI principles of communication and self-care was especially helpful in this regard (Table 1 and Table 2).

Table 1. Moderator feedback on CAN Quit (Connecting Alaska Native People to Quit Smoking) beta test.

Question and moderator feedback	Changes made
What were the most challenging parts of moderating the CAN Quit group?	
Maintaining communication among participants on multiple threads at one time. It can be easy to miss someone's response if you are not diligent.	Study team monitors group and notifies moderators if they miss a post
There were participants who shared personal experiences (eg, death of a child and struggles with other substances) with the group, and it was challenging to figure out how to best respond to these comments. There was also a participant who used our post to sell items on a personal FB[a] page.	• Consultation with the trainer (CY) before crafting a response • Determine as a group the best way to handle more sensitive topics • Added an additional refresher training in MI[b] principles (KR) and practices before RCT[c]
Tracking each participant's responses so you can easily reference prior conversations. This is important when building rapport with the participant.	Moderators maintain a spreadsheet tracking all members with notes to help them remember details and respond accordingly
Connecting one person to another based on their personal experiences and actually getting a response or any dialogue between the two.	Feedback or reinforcement from the trainer (CY) on different tactics for connecting participants (ie, tagging, private messaging, and asking questions specific to interests or strengths of participants).
What advice would you give to someone who was about to moderate a group like CAN Quit?	
Keep a log of your communication with each participant from when the group starts. This way you can easily reference this spreadsheet to see what topics have been discussed with this participant. This can be especially helpful in groups with more than one moderator.	N/A[d]
Be thoughtful in your responses; sometimes it can be difficult to not sound robotic over the computer. I think this is important in developing a connection with each participant and building trust. The MI tips we learned were very helpful in making sure we are professional but personable when moderating.	N/A
What did you think you did best and what do you think you need to work on?	
We did a good job of posting new content with an interesting prompt every few days which helped the group stay active.	N/A
We did a good job of refining our responses to each participant and knowing when to refer the Quitline. We do not want to refer the Quitline so often that we sound robotic but we want to encourage participants to use evidence-based smoking cessation programs.	Use MI skills training and created additional posts that discuss the Quitline and Tribal quitting resources
We are working on generating quicker responses	Moderators share the responsibility of monitoring the Quitline

[a]FB: Facebook.

[b]MI: motivational interviewing.

[c]RCT: randomized controlled trial.

[d]N/A: not applicable.

Table 2. Beta tester feedback on CAN Quit (Connecting Alaska Native People to Quit Smoking) beta test.

Improvement areas (themes)	Representative quotes	Solutions implemented by team
Clarify group purpose	• "Have mental or behavioral health programs ready" • "I felt unsure if I was supposed to use the site to access the quit helpline through FB[a], because my region doesn't use the Quitline for assistance in quitting anymore"	• Additional welcome posts created that clarify group purpose as well as the Quitline and Tribal Quitting resources • Moderators incorporated feedback to reinforce FB page purpose and the quitting resources
Provide incentives or games	• "It would be great to get the rural communities involved and offer some sort of Photo, Video or Art contest that included a drawing, a photo or video that helps people quit smoking or talk about the negative effects of smoking. It would draw more people in and get people to participate in the efforts toward Tobacco Prevention and keep them busy toward something positive. If we could include a prize for 1st, 2nd and 3rd also!" • "Maybe some fun trivia games to get people wanting to look at the page." • "Do more instead of just post posters. I mean we all can brainstorm in a different way and play along as we go?"	• Team implemented "polls" where participants could vote on what values are important to them and write in additional values • Will add incentives for future groups
How to improve posts	• "Could use less medical type photos. More of our own pictures to help support each other" • "I think more should be put into it. More involvements" • "More videos that include rural areas would be a great idea" • "Posts being more clear about regional resources would be nice" • "Make more posts"	• More rural posts will be used in the randomized controlled trial • More posts relating to Alaska Native values were created using Alaska Native imagery (eg, people, families, communities, and activities) • New post and discussion prompt posted to the group every 2 days • Combined posts and video links to increase participant engagement • Creation of interactive posts like the FB poll • Developed a moderator spreadsheet to track posts to ensure full variety of topics and types of posts (eg, urban, rural, photo or text, and video) will be addressed
Moderators make group more interactive	• "Maybe encourage more people to post more often to gain connection." • "Find ways to make it more interactive."	• Moderators will tag participants to increase participation on individual posts • Moderators individually welcome all new members on the welcome post and encourage them to interact. • Prompted new members to introduce themselves and their reasons for wanting to quit. • Moderators connect members to interact who express similar experiences or concerns • Provided additional training to moderators that focused on health communication and ways of increasing interaction • Developed follow-up questions that may be asked on the original post to encourage discussion

[a]FB: Facebook.

Conclusions, Discussion, and Overall Recommendations

This paper shares the processes involved with developing and beta-testing a social media group as an intervention to promote use of evidence-based resources for quitting smoking using peer support. We share specific approaches, strategies, and methods for stimulating participant engagement in the Facebook group and emphasize the importance of moderator training and creating a content library. During training, trainers and moderators shared expertise to help build cultural competency and web-based communication skills that would build community and promote behavior change. Cultural tailoring is a critical component of the prototype. The research team used content linked to core Alaska Native values and/or Alaska Native imagery when possible, and Alaska Native moderators used their personal experience to incorporate cultural values and content into their written posts and responses. The purpose of these phases was to adapt the intervention and prepare moderators to deliver the intervention for the subsequent phase, which was a pilot RCT.

The results of the beta test were encouraging with participants expressing a high degree of usability of the CAN Quit Facebook group based on the Social Media Usability Measure, and there was a signal of potential effectiveness, with 2 out of 10 participants using the Quitline and reporting quitting smoking.

Qualitative feedback was helpful in directing the research team to specific improvements needed to make the Facebook group more successful or usable; for example, providing additional training to moderators to promote interactions among participants, support conversations that may contain emotional content, and add more welcome posts to the group page to better clarify the group's purpose: use of evidence-based quitting resources. Given that the group is on Facebook, which is popular in Alaska Native communities [16], most users were not new to the technology, which made it easier to avoid technical challenges in participating. We noticed that no video content was posted by moderators and some other content was posted repeatedly, whereas other content was not used. The selection of content was based on the activity of their group members at the discretion of the moderators. However, based on this observation, we determined that it would be helpful to the moderators as they select content, to develop a spreadsheet to track posts already used from the content library.

Strategies to promote engagement between participants rather than between participants and the moderator are the primary areas for improvement. Strategies suggested by our community management expert (CY) included tagging, private messaging, pinning posts so they show up at the top of the group feed, and drafting new follow-up questions to add to posts to further stimulate discussion. Several strengths were noted during the beta-testing phase. Moderators were able to model empathy and support, which included their adept use of reflective listening skills and personal experiences as tobacco counselors. Both moderators had received TTS training and were of Alaska Native heritage, which made it easier for them to quickly connect to participants through a mutual appreciation of core Alaska Native values and their TTS background.

Our approach has several strengths. The development of our training and content library was iterative and theory based. We used social media and health communication experts to conduct training grounded in situational learning. We used a combination of face-to-face and web-based formats and provided a scenario-based approach for moderators to practice their skills. All participants (moderators and trainers) in the training sessions served as experts to each other such that all participants were able to acquire new skills. We gathered feedback from the research team and our beta testers on how to improve the CAN Quit prototype. This feedback resulted in practical solutions that will be embedded in the RCT, which is beneficial for developing an intervention in phases (iteratively), as recommended by the National Institutes of Health and CDC guidelines [47,48]. Most importantly, our approach was created based on the need to improve use of evidence-based tobacco cessation resources specifically for Alaska Native people, a high-risk group for tobacco-related morbidity and mortality. This approach provided a systematic process to tailor or adapt content and training to appeal specifically to Alaska Native people living in Alaska.

Our approach also has some limitations. Implementing the prototype on Facebook results in a lack of control on some aspects of the intervention design, such as the order in which posts were introduced, which required moderator flexibility in responding to changes accordingly. Despite this limitation, given that Facebook is commonly used and familiar, less on-boarding of participants is required to participate in the intervention. In addition, a sample size of 10 was the minimum recommended for a group social media intervention; it is likely that engagement would be easier to achieve with a larger group size, as will be used for the RCT. Stimulating group interaction during the beta-testing phase presented a challenge to the moderators given the small number of participants and the short amount of time they participated. However, this challenge resulted in stronger engagement skills among moderators, which will aid in the RCT phase. Meanwhile, consistent with and building on the work of Pagoto et al [29], we came up with ten recommendations of lessons learned that could help guide others interested in developing similar social media intervention prototypes: (1) select moderators who can relate to your group; (2) build on existing skills of your moderators; (3) build community engagement; (4) use existing resources for content; (5) adapt content for cultural fit; (6) design content that highlights your goals; (7) include a wide variety of content and organize it into an easily searched library; (8) balance content with conversation; (9) support moderators to handle emotional communication; and (10) groom group participants as moderators. Table 3 outlines these recommendations in detail.

In summary, providing practical information on how we developed and evaluated our intervention prototype demonstrates how a beta-testing phase is essential in identifying how the platform, content, and moderation can be improved before rolling out a large-scale intervention. We learned that a social media group intervention is most likely to succeed if it balances well-developed culturally congruent content with strategies to engage Alaska Native participants in social media conversations facilitated by moderators. Moderators need to be provided with the necessary skills and methods for promoting web-based community engagement with and among group members and in encouraging behavior change. In our case, having Alaska Native moderators who were themselves Alaska Native and already experienced the challenges associated with smoking cessation was extremely helpful in both developing the training plan and promoting engagement. Future studies could explore using former intervention participants transitioning to a moderator role to enhance sustainability, growth, and feasibility of the intervention. Our next step will be to evaluate the prototype via a pilot RCT.

Table 3. Recommendations for creating a successful social media group to promote use of evidence-based smoking cessation resources for Alaska Native people.

Recommendation	Description
Select moderators who can relate to your group	We chose moderators with expertise in smoking cessation (TTS[a]) who also shared similar Alaska Native ancestry and cultural values. This approach is effective with building rapport among group participants and moderators.
Build upon existing skills of your moderators	Beyond training as TTS, our moderators received 20 additional hours of training in strategic web-based community management and health behavior change. We recommend training be scenario-based, experiential, and use actual posts for moderating practice.
Build community engagement	Our moderators shared that getting to know group participants made it easier to respond to their posts, reinforce quitting messages, and provide support. Train moderators to react, tag, and instant message participants to respond when they post. They should acknowledge active participants while also checking in with inactive participants, inviting them back to the group.
Use existing resources for content	We evaluated and adapted existing, evidence-based content already created by the CDC[b] and ANTHC[c], with their permission. This was both cost and time effective. We recommend using or refining existing high-quality content whenever possible, seeking necessary permissions.
Adapt content for cultural fit	We used an iterative theory-based process of qualitative and quantitative feedback from both participants and stakeholders to adapt our content and be responsive to the evolution of the participants and the group. We recommend using a conceptual framework with community and stakeholder input.
Design content that highlights your goals	Our goal was to promote use of evidence-based resources. We found it important to include links to the Alaska Quitline or Tribal cessation resources on all content. We suggest including a call to action with each post.
Include a wide variety of content and organize it into an easily searched library	It is difficult to predict how participants will communicate and what will motivate them to engage in the group. Having an organized library with a variety of content allows moderators flexibility to select content that is responsive to an ongoing conversation or that meets the needs or interests of group participants to maintain engagement.
Balance content with conversation	We used content that fit concerns and conversations happening within the group. We believe the key to success lies in fostering group engagement through conversation with moderators and with each other. It is necessary to create a community of talkers, not readers.
Support moderators to handle emotional communication	We provided 20 hours of MI[d] and web-based community management training to prepare moderators to support group participants who share sensitive information. We recommend training also include a plan for moderators to seek additional support.
Groom group participants to be moderators	Once our group was underway, we saw some participants were very engaged, not only with moderators but also with other group participants. Consider inviting and training former group participants to volunteer as peer mentors or moderators.

[a]TTS: trained tobacco specialist.

[b]CDC: Centers for Disease Control and Prevention.

[c]ANTHC: Alaska Native Tribal Health Consortium.

[d]MI: motivational interviewing.

Acknowledgments

Authors thank the Alaska Native Tribal Health Consortium Research Consultation Committee for providing guidance on the development and implementation of this study. They also appreciate the contributions of Selma Oskolkoff-Simon, Fiona Brosnan, Michael Doyle, and Alaska Native Tribal Health Consortium Marketing and Communication. The authors are grateful for the technical assistance provided by the Health Communications Branch of the Centers for Disease Control and Prevention. In addition, the authors would like to acknowledge the Mayo Clinic Survey Research Center for their assistance with the survey design and data collection. They also thank Michelle Pearson and Kimberly Kinnoin for their assistance with the manuscript.

This research was supported by the National Institute on Drug Abuse of the National Institutes of Health under award number R34DA046008 (Patten). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

We thank the staff at the Office of Smoking and Health at the Centers for Disease Control and Prevention, who arranged access to the image used in this study.

Conflicts of Interest

Unrelated to this project, JJP has provided consultation to pharmaceutical and technology companies that make medications and other treatments for quitting smoking and has received funding from Facebook for planning evaluation of a mobile health intervention. JJP has also served as an expert witness in lawsuits against tobacco companies. All other authors declare no conflicts of interest.

References

1. Alaska Tobacco Facts 2018. Alaska Department of Health and Social Services (AKDHSS). 2018. URL: http://dhss.alaska.gov/dph/Chronic/Documents/Tobacco/PDF/2018_AKTobaccoFacts.pdf [accessed 2019-01-26]
2. Alaska Tobacco Facts 2019. Alaska Department of Health and Social Services. 2019. URL: https://dhss.alaska.gov/dph/Chronic/Documents/Tobacco/PDF/2019_AKTobaccoFacts.pdf [accessed 2022-01-11]
3. Jamal A, Phillips E, Gentzke AS, Homa DM, Babb SD, King BA, et al. Current cigarette smoking among adults - United States, 2016. MMWR Morb Mortal Wkly Rep 2018 Jan 19;67(2):53-59 [FREE Full text] [doi: 10.15585/mmwr.mm6702a1] [Medline: 29346338]
4. Cobb NK, Graham AL, Byron MJ, Niaura RS, Abrams DB, Workshop P. Online social networks and smoking cessation: a scientific research agenda. J Med Internet Res 2011;13(4):e119 [FREE Full text] [doi: 10.2196/jmir.1911] [Medline: 22182518]
5. Civljak M, Sheikh A, Stead LF, Car J. Internet-based interventions for smoking cessation. Cochrane Database Syst Rev 2010(9):CD007078. [doi: 10.1002/14651858.CD007078.pub3] [Medline: 20824856]
6. Hutton HE, Wilson LM, Apelberg BJ, Tang EA, Odelola O, Bass EB, et al. A systematic review of randomized controlled trials: web-based interventions for smoking cessation among adolescents, college students, and adults. Nicotine Tob Res 2011 Apr;13(4):227-238. [doi: 10.1093/ntr/ntq252] [Medline: 21350042]
7. Bernhardt JM, Mays D, Kreuter MW. Dissemination 2.0: closing the gap between knowledge and practice with new media and marketing. J Health Commun 2011;16 Suppl 1:32-44. [doi: 10.1080/10810730.2011.593608] [Medline: 21843094]
8. Soto MV, Balls-Berry JE, Bishop SG, Aase LA, Timimi FK, Montori VM, et al. Use of web 2.0 social media platforms to promote community-engaged research dialogs: a preliminary program evaluation. JMIR Res Protoc 2016;5(3):e183 [FREE Full text] [doi: 10.2196/resprot.4808] [Medline: 27613231]
9. Myung S, McDonnell DD, Kazinets G, Seo HG, Moskowitz JM. Effects of web- and computer-based smoking cessation programs: meta-analysis of randomized controlled trials. Arch Intern Med 2009 May 25;169(10):929-937. [doi: 10.1001/archinternmed.2009.109] [Medline: 19468084]
10. Mohatt G, McDiarmid G, Montoya V. Societies, families,change: the Alaskan example. In: Manson S, Dinges N, editors. Behavioral Health Issues Among American Indians and Alaska Natives. Aurora, CO: Centers for American Indian & Alaska Native Health; 1988:325-365.
11. Teufel-Shone NI, Tippens JA, McCrary HC, Ehiri JE, Sanderson PR. Resilience in American Indian and Alaska native public health: an underexplored framework. Am J Health Promot 2018 Feb;32(2):274-281 [FREE Full text] [doi: 10.1177/0890117116664708] [Medline: 27577566]
12. Wolsko C, Lardon C, Hopkins S, Ruppert E. Conceptions of wellness among the Yup'ik of the Yukon-Kuskokwim Delta: the vitality of social and natural connection. Ethn Health 2006 Nov;11(4):345-363. [doi: 10.1080/13557850600824005] [Medline: 17060033]
13. Barnhardt R, Kawagley AO. Indigenous knowledge systems and Alaska native ways of knowing. Anthropol Edu Q 2008 Jan 08;36(1):8-23 [FREE Full text] [doi: 10.1525/aeq.2005.36.1.008]
14. Carmack E, Macdonald R. Water and ice-related phenomena in the coastal region of the Beaufort sea: some parallels between native experience and western science. ARCTIC 2009 Aug 25;61(3):265-280. [doi: 10.14430/arctic24]
15. Eighth broadband progress report. Federal Communications Commission (FCC). URL: https://www.fcc.gov/reports-research/reports/broadband-progress-reports/eighth-broadband-progress-report [accessed 2017-01-17]
16. Greenwood S, Perrin A, Duggan M. Social media update. Pew Research Center. 2016. URL: http://www.pewinternet.org/2016/11/11/social-media-update-2016/ [accessed 2017-03-02]
17. Hudson HE. Digital diversity: broadband and indigenous populations in Alaska. J Inform Policy 2011;1:378. [doi: 10.5325/jinfopoli.1.2011.0378]
18. Hudson H. Toward universal broadband in rural Alaska: a report of the institute of social and economic research. Part 1: an analysis of internet use in southwest alaska; Part 2: literature review. University of Alaska Anchorage, Anchorage, AK. 2012. URL: https://www.alaska.edu/files/oit/bbtaskforce/docs/2012-11-ISER_Terra-SW_Study.pdf [accessed 2017-03-18]
19. Reese N. Alaska's broadband: stats and figures. Broadband Now. 2017. URL: http://broadbandnow.com/Alaska [accessed 2017-02-01]
20. Power J, Bersamin A. Exploring the potential for technology-based nutrition education for Alaska Native WIC recipients in rural Southwest Alaska. FASEB J 2016;30(1 Supplement):lb430 [FREE Full text] [doi: 10.1096/fasebj.30.1_supplement.lb430]

21. Passey M, Bonevski B. The importance of tobacco research focusing on marginalized groups. Addiction 2014 Jul;109(7):1049-1051. [doi: 10.1111/add.12548] [Medline: 24758261]
22. Sinicrope PS, Koller KR, Prochaska JJ, Hughes CA, Bock MJ, Decker PA, et al. Social media intervention to promote smoking treatment utilization and cessation among Alaska native people who smoke: protocol for the connecting Alaska native people to quit smoking (Can Quit) pilot study. JMIR Res Protoc 2019 Nov 22;8(11):e15155 [FREE Full text] [doi: 10.2196/15155] [Medline: 31755867]
23. Merculieff ZT, Koller KR, Sinicrope PS, Hughes CA, Bock MJ, Decker PA, et al. Developing a social media intervention to connect Alaska native people who smoke with resources and support to quit smoking: the connecting Alaska native quit study. Nicotine Tob Res 2021 May 24;23(6):1002-1009. [doi: 10.1093/ntr/ntaa253] [Medline: 33674856]
24. Dohan D, Garrett SB, Rendle KA, Halley M, Abramson C. The importance of integrating narrative into health care decision making. Health Aff (Millwood) 2016 Apr;35(4):720-725. [doi: 10.1377/hlthaff.2015.1373] [Medline: 27044974]
25. Gubrium A. Digital storytelling: an emergent method for health promotion research and practice. Health Promot Pract 2009 Apr;10(2):186-191. [doi: 10.1177/1524839909332600] [Medline: 19372280]
26. Kreuter MW, Green MC, Cappella JN, Slater MD, Wise ME, Storey D, et al. Narrative communication in cancer prevention and control: a framework to guide research and application. Ann Behav Med 2007 Jun;33(3):221-235. [doi: 10.1080/08836610701357922] [Medline: 17600449]
27. McAfee T, Davis KC, Alexander RL, Pechacek TF, Bunnell R. Effect of the first federally funded US antismoking national media campaign. Lancet 2013 Dec 14;382(9909):2003-2011. [doi: 10.1016/S0140-6736(13)61686-4] [Medline: 24029166]
28. Davis R, Resnicow K. The cultural variance framework for tailoring health messages. In: Cho H, editor. Health Communication Message Design: Theory and Practice. Thousand Oaks, CA: Sage Publications, Inc; 2012:115-135.
29. Pagoto S, Waring ME, May CN, Ding EY, Kunz WH, Hayes R, et al. Adapting behavioral interventions for social media delivery. J Med Internet Res 2016;18(1):e24 [FREE Full text] [doi: 10.2196/jmir.5086] [Medline: 26825969]
30. Millington R. Introducing the map - a proven process for developing successful online communities. FeverBee. 2012. URL: https://www.feverbee.com/introducing-the-map-a-proven-process-for-developing-successful-online-communities/ [accessed 2022-01-11]
31. Wenger E, McDermott R, Snyder W. Cultivating Communities of Practice. Boston: Harvard Business Review Press; 2002.
32. Young C. Community management that works: how to build and sustain a thriving online health community. J Med Internet Res 2013;15(6):e119 [FREE Full text] [doi: 10.2196/jmir.2501] [Medline: 23759312]
33. Johnson L. Social network vs. online community: what is the difference? Social Media Today. 2014. URL: https://www.socialmediatoday.com/content/social-network-vs-online-community-what-difference [accessed 2022-01-11]
34. Lave J, Wenger E. Situated Learning: Legitimate Peripheral Participation. Cambridge: Cambridge University Press; 1991.
35. Harrison JD, Auerbach AD, Anderson W, Fagan M, Carnie M, Hanson C, et al. Patient stakeholder engagement in research: a narrative review to describe foundational principles and best practice activities. Health Expect 2019 Jun;22(3):307-316 [FREE Full text] [doi: 10.1111/hex.12873] [Medline: 30761699]
36. Mayo clinic connect. Mayo Clinic. URL: https://connect.mayoclinic.org/ [accessed 2022-01-11]
37. Skousen T, Safadi H, Young C, Karahanna E, Safadi S, Chebib F. Successful moderation in online patient communities: inductive case study. J Med Internet Res 2020 Mar 17;22(3):e15983 [FREE Full text] [doi: 10.2196/15983] [Medline: 32181743]
38. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [doi: 10.1037//0003-066x.55.1.68] [Medline: 11392867]
39. Beck AT. Cognitive therapy: nature and relation to behavior therapy. Behav Ther 1970 May;1(2):184-200. [doi: 10.1016/s0005-7894(70)80030-2]
40. Ellis A. Reason and Emotion in Psychotherapy. New York: Lyle Stuart; 1962.
41. Resnicow K, Baranowski T, Ahluwalia JS, Braithwaite RL. Cultural sensitivity in public health: defined and demystified. Ethn Dis 1999;9(1):10-21. [Medline: 10355471]
42. Heckman CJ, Egleston BL, Hofmann MT. Efficacy of motivational interviewing for smoking cessation: a systematic review and meta-analysis. Tob Control 2010 Oct;19(5):410-416 [FREE Full text] [doi: 10.1136/tc.2009.033175] [Medline: 20675688]
43. Ramo DE, Thrul J, Chavez K, Delucchi KL, Prochaska JJ. Feasibility and quit rates of the tobacco status project: a Facebook smoking cessation intervention for young adults. J Med Internet Res 2015;17(12):e291 [FREE Full text] [doi: 10.2196/jmir.5209] [Medline: 26721211]
44. Ramo DE, Thrul J, Delucchi KL, Ling PM, Hall SM, Prochaska JJ. The Tobacco Status Project (TSP): study protocol for a randomized controlled trial of a Facebook smoking cessation intervention for young adults. BMC Public Health 2015 Sep 15;15:897 [FREE Full text] [doi: 10.1186/s12889-015-2217-0] [Medline: 26374203]
45. Lund A. Measuring usability with the USE Questionnaire. Usability Interface. 2001. URL: https://www.researchgate.net/publication/230786746_Measuring_Usability_with_the_USE_Questionnaire [accessed 2022-01-11]
46. Krippendorff K. Content Analysis: An Introduction to Its Methodology. 3rd Ed. Thousand Oaks, CA: SAGE Publications; 2012.

47. CDCSocialMediaWorks: developing a strategy for using social media. Centers for Disease control and Prevention (CDC). 2016. URL: https://www.cdc.gov/socialmedia/Tools/guidelines/pdf/SocialMediaToolkit_BM.pdf?source=govdelivery [accessed 2017-05-08]
48. National Cancer Institute. Making health communictions programs work. US Department of Health and Human Services. 2004. URL: https://www.cancer.gov/publications/health-communication/pink-book.pdf [accessed 2017-06-01]

Abbreviations

ANTHC: Alaska Native Tribal Health Consortium
CAN Quit: Connecting Alaska Native People to Quit Smoking
CDC: Centers for Disease Control and Prevention
MI: motivational interviewing
RCT: randomized controlled trial
TTS: trained tobacco specialist

Edited by A Mavragani; submitted 12.03.21; peer-reviewed by H Safadi, V Johansson; comments to author 28.04.21; revised version received 21.06.21; accepted 01.12.21; published 17.02.22.

Please cite as:
Sinicrope PS, Young CD, Resnicow K, Merritt ZT, McConnell CR, Hughes CA, Koller KR, Bock MJ, Decker PA, Flanagan CA, Meade CD, Thomas TK, Prochaska JJ, Patten CA
Lessons Learned From Beta-Testing a Facebook Group Prototype to Promote Treatment Use in the "Connecting Alaska Native People to Quit Smoking" (CAN Quit) Study
J Med Internet Res 2022;24(2):e28704
URL: https://www.jmir.org/2022/2/e28704
doi:10.2196/28704
PMID:35175208

Viewpoint

The 21st Century Cures Act and Multiuser Electronic Health Record Access: Potential Pitfalls of Information Release

Simone Arvisais-Anhalt[1,2], MD; May Lau[3], MD, MPH; Christoph U Lehmann[4,5], MD; A Jay Holmgren[6], PhD; Richard J Medford[7], MD; Charina M Ramirez[8], MD; Clifford N Chen[9], MD

[1]Department of Hospital Medicine, University of California San Francisco, San Francisco, CA, United States

[2]Department of Laboratory Medicine, University of California San Francisco, San Francisco, CA, United States

[3]Division of Developmental and Behavioral Pediatrics, Department of Pediatrics, University of Texas Southwestern Medical Center and Children's Medical Center Dallas, Dallas, TX, United States

[4]Department of Pediatrics, University of Texas Southwestern Medical Center, Dallas, TX, United States

[5]Department of Data Sciences and Bioinformatics, University of Texas Southwestern Medical Center, Dallas, TX, United States

[6]Department of Medicine, Center for Clinical Informatics and Improvement Research, University of California San Francisco, San Francisco, CA, United States

[7]Division of Infectious Disease, Department of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, TX, United States

[8]Division of Pediatric Gastroenterology, Hepatology, and Nutrition, Department of Pediatrics, University of Texas Southwestern Medical Center and Children's Medical Center Dallas, Dallas, TX, United States

[9]Division of Hospital Medicine, Department of Pediatrics, University of Texas Southwestern Medical Center and Children's Medical Center Dallas, Dallas, TX, United States

Corresponding Author:
Simone Arvisais-Anhalt, MD
Department of Laboratory Medicine
University of California San Francisco
400 Parnassus Avenue
San Francisco, CA, 94143
United States
Phone: 1 415 476 1000
Email: simone.arvisais-anhalt@ucsf.edu

Abstract

Although the Office of The National Coordinator for Health Information Technology's (ONC) Information Blocking Provision in the Cures Act Final Rule is an important step forward in providing patients free and unfettered access to their electronic health information (EHI), in the contexts of multiuser electronic health record (EHR) access and proxy access, concerns on the potential for harm in adolescent care contexts exist. We describe how the provision could erode patients' (both adolescent and older patients alike) trust and willingness to seek care. The rule's preventing harm exception does not apply to situations where the patient is a minor and the health care provider wishes to restrict a parent's or guardian's access to the minor's EHI to avoid violating the minor's confidentiality and potentially harming patient-clinician trust. This may violate previously developed government principles in the design and implementation of EHRs for pediatric care. Creating legally acceptable workarounds by means such as duplicate "shadow charting" will be burdensome (and prohibitive) for health care providers. Under the privacy exception, patients have the opportunity to request information to not be shared; however, depending on institutional practices, providers and patients may have limited awareness of this exception. Notably, the privacy exception states that providers cannot "improperly encourage or induce a patient's request to block information." Fearing being found in violation of the information blocking provisions, providers may feel that they are unable to guide patients navigating the release of their EHI in the multiuser or proxy access setting. ONC should provide more detailed guidance on their website and targeted outreach to providers and their specialty organizations that care for adolescents and other individuals affected by the Cures Act, and researchers should carefully monitor charting habits in these multiuser or proxy access situations.

(J Med Internet Res 2022;24(2):e34085) doi:10.2196/34085

KEYWORDS

21st Century Cures Act; Open Notes; Information Blocking; multiuser EHR access; proxy EHR access; adolescent Health; health IT Policy; information technology; cures act; electronic health record; electronic health information; health information; patient care

Introduction

"*Primum non nocere*" ("*First, do no harm*") or nonmaleficence is a fundamental principle taught to every health care provider. It suggests that before applying any medical intervention, one needs to consider the potential negative effects on the patient. In this piece, we examine the potential for patient harm by the Office of The National Coordinator for Health Information Technology's (ONC) Information Blocking Provision in the Cures Act Final Rule and the additional burden that health care providers, those who provide patient care and provide documentation in the electronic health record, will now face when documenting sensitive information.

On December 13, 2016, the 21st Century Cures Act (hereinafter referred to as the "Cures Act") was signed into law with the intent to "accelerate the discovery, development, and delivery of 21st century cures, and for other purposes" [1]. The Act defined electronic health record (EHR) interoperability, addressed health information technology certification requirements, and prohibited information blocking—the practice that prevents or interferes with those with permission to access electronic health information (EHI) [2]. As the federal entity coordinating efforts to implement health information technology and exchange EHI, ONC, a division within the US Department of Health and Human Services (HHS) [3], developed the Cures Act Final Rule to direct the implementation of the Cures Act legislation [4].

The ONC Cures Act Final Rule

The stated goal of the ONC Cures Act Final Rule is to empower patients to interact "with their health record in a modern health IT economy" [4]. ONC postulated that "putting patients in charge of their health record is a key piece of patient control in healthcare and patient control is at the center of HHS's work towards a value-based healthcare system." The Cures Act Final Rule also encourages innovations in health care technology and hopes to deliver the following:

- Transparency on cost and outcomes of care
- Competitive options in obtaining medical care
- Convenient access to medical records using smartphone apps
- Innovation and choice for patients, physicians, hospitals, payers, and employers through an app-based economy [4]

The Cures Act Final Rule promotes interoperability across EHR vendors through the adoption of data exchange standards and calls upon the health care information technology (IT) industry to adopt standardized application programming interfaces through specified Conditions of Certification. Additionally, ONC aims to increase patients' access to their EHI through minimizing measures that block patient access to information [5-7].

The Information Blocking Provision

The Information Blocking Provision of the Cures Act Final Rule mandates that patients have unfettered, free access to their EHI, and provides clear requirements for compliance by health care providers, institutions, health information exchanges, and EHR vendors [8].

The spirit of the Information Blocking Provision is similar to that of the OpenNotes movement, which over the past decade has been adopted by several health care institutions across the United States, Canada, and Sweden and provides patients with near immediate and full access to their EHI [9,10]. The Information Blocking Provision requires that patients have access to parts of their EHI defined by the United States Core Data for Interoperability (Figure 1) by April 5, 2021, with eventual expansion to all EHI by October 6, 2022 [11-13]. Of note, patients have had the right to access their medical record since the implementation of the Health Insurance Portability and Accountability Act (HIPAA). The Cures Act final rule does not increase the type of health information that patients and families can access, it only facilitates automatic release via patient portals and easier access electronically.

The Information Blocking Provision includes a means to report violations and enforcement options. ONC encourages anyone who experiences or observes information blocking by any health care provider, health IT developer, certified health IT, health information network, or information exchange to share their concerns through an information blocking portal on ONC's website [11]. Health IT developers, health information networks, and health information exchanges can be subject to civil monetary penalties of up to US $1,000,000 per violation [14]. Health care providers found to have committed information blocking will also be subjected to penalties that are to be determined [14].

Figure 1. Elements of the United States Core Data for Interoperability [11].

Exceptions to the Information Blocking Provision

The Information Blocking Provision defines eight exceptions that do not constitute information blocking [15]. The preventing harm exception and the privacy exception are applicable to the documenting health care provider.

The preventing harm exception stipulates that provided certain conditions are met, a health care provider can prevent the access to a patient's EHI if it is "reasonable and necessary to prevent harm to a patient or another person" [15]. Key conditions include that the health care provider must reasonably believe that preventing access to a patient's EHI will significantly reduce a risk of substantial harm, and that the interference is no broader than necessary. The patient has the right to request a review of an individualized determination of risk of harm [16].

According to ONC's guidance in the "Information Blocking Frequently Asked Questions," the "Preventing Harm" exception does not apply to situations where the patient is a minor and the health care provider wishes to restrict a parent or legal representative's access to the minor's EHI to avoid violating the minor's confidentiality and destroying the trust between the youth and the health care provider [12]. This lack of applicability in the case of adolescent confidentiality stands in tension with principles outlined by experts in the design and implementation of EHRs [17], endorsed by the American Academy of Pediatrics and Society of Adolescent Health And Medicine [18,19]. The concern over the implication of the Cures Act on adolescent confidentiality has been noted in the literature [20]. The premise underlying confidential care encourages adolescents to communicate with health care providers about sensitive topics such as sexual and reproductive health and substance abuse without the fear that their parents or guardians will have access to this information. Confidentiality on certain health care problems facilitates obtaining medical care that adolescents might forgo if information were shared with others. In the context of providing confidential care, the Cures Act's broad focus on patient EHI access may cause a trade-off with patient–provider relationships, trust, and nonphysical types of harm. The text of the Final Rule specifically states that the desire to maintain confidentiality and to protect patient–provider relationships is insufficient to prevent the release of sensitive information. In certain multiuser access cases, this may erode the patient's control over his/her information instead of increasing control. It is worth noting that HIPAA and the Cures act defer to the state laws that grant adolescents the ability to consent for certain conditions. While it is challenging to keep track of each state's individual and varied confidentiality laws that result in 56 (one for each state and territory) different legal requirements for users of pediatric EHRs, these laws do provide clear legal backing to protect adolescent confidentiality.

The other exception to the Cures Act Final Rule's patient access provision immediately relevant to health care providers is the privacy exception. Under this exception, interfering with access to EHI is deemed not to be information blocking when the intent is to protect the patient's privacy. These exceptions, listed in Textbox 1, are included by ONC to comply with HIPAA and other state privacy regulations and allow patients the opportunity to request information not be shared. Depending upon institutional practices, providers and patients may have limited awareness of this exception. Notably, the privacy exception states that providers cannot "improperly encourage or induce a patient's request to block information" [21]. This stipulation affects a provider's ability to guide patients for fear of being found in violation of the Information Blocking Provision and fined. Institutional policies and procedures will affect the implementation and management of this exception. Providers may not be aware of the procedure for patients to request their information not be shared during an encounter. Depending upon when the patient places the request (eg, before, during, or after an encounter), the institution may not be able to fulfill the request in a timely manner relative to the immediacy of information being released. Additionally, the privacy exception does not clearly describe if and how a patient can block individual pieces of data (data segmentation) instead of all data. The exception only describes how patients can request to block access and can request to regain access. The interpretation and implementation of this exception is left to the institution and provider and, given the complex nature of the exception, necessitates deference to informatics expertise and legal resources with experience in state and federal privacy laws and statutes for interpretation and use.

Textbox 1. Privacy exceptions to information blocking.

Exceptions:

- More stringent state or federal preconditions to exchange is not met
- Information technology developer is not covered by the Health Insurance Portability and Accountability Act Privacy Rule
- Inability to validate a requester's right to access
- The individual requests the information not to be shared

Limitations to the Information Blocking Provision: Multiuser or Proxy Access

Overview

The OpenNotes initiative has been shown to potentially increase patient activation, engagement, satisfaction, trust, and safety, and to improve the patient–physician relationship [22-25]. However, concern exists that the Information Blocking Provision will result in damaging breaches of confidentiality for cohorts of patients when parents or legal representatives are provided multiuser or proxy access to EHI [26,27]. In circumstances where EHI is made available within a web-based portal with multiuser or proxy access, the information could compromise the confidentiality of the patient, parent, or legal representatives and damage the relationship between the health care provider, patient, parent, or legal representative. The breach of confidentiality may occur bidirectionally as a caregiver may share information with a provider, which could be shared back with the patient. One recent study highlighted another area of concern: when guardians access an adolescent patient's portal account. The study revealed that the estimated prevalence of guardian access could be as high as 76% of adolescent accounts and also showed a relatively low rate of proxy account creation [28]. When adolescents had their own portal account, proxy accounts for adolescent patients were created in only 0.3%-10% of cases [28]. The reality that many portal accounts are used and managed by guardians must be taken into consideration for adolescent patients who, in the context of their care setting, may lack the autonomy to prevent their guardians from accessing their personal patient portals.

Pediatric and Adolescent Patients

Prior to the Information Blocking Provision, pediatric institutions participating in the OpenNotes movement had addressed the concern for violating confidentiality and damaging relationships by blocking all clinical notes from several clinics including adolescent, gynecology, psychiatry, substance abuse, and the child protection team [29]. Although the Cures Act Final Rule explicitly states that maintaining confidentiality and protecting relationships is not sufficient to prevent the release of sensitive information, the effects of releasing this information on patients, their parents or legal representatives, and the patient–provider relationship cannot be underestimated and are concerning to adolescent medicine providers and other health care providers who care for youth [30].

The Adolescent–Health Care Provider Relationship

There are many situations that do not fit the "Preventing Harm" exception where adolescent patients may be adversely affected when their private information is accessed by others (Table 1). For example, an adolescent female with concerns for a sexually transmitted infection (STI) such as *Neisseria gonorrhoeae* or *Chlamydia trachomatis* may avoid seeking medical care to avoid repercussions or stigma if she knew her parents would have access to this information. This untreated STI could progress to pelvic inflammatory disease, a more serious infection, which may require hospitalization and intravenous antibiotic administration and could affect future fertility. Prior research has shown that 59% of surveyed females younger than 18 years would "stop using all sexual healthcare services, delay testing or treatment for HIV or other STDs, or discontinue use of specific (but not all) sexual healthcare services if their parents were informed they were seeking prescribed contraceptives" [31]. The concern for loss of confidentiality extends to other sensitive topics including mental health, substance use, gender identity, and sexual orientation and may conflict with federal and state laws.

Table 1. Hypothetical scenarios for potential harm related to either lack of clarity of the laws, technical limitations regarding the release of electronic health information, or a combination of both.

At risk for harm	Third party receiving information	Domain	Mode of disclosure	Consequence
Patient	Parents or guardian	Mental health	Patient portal	Avoiding care or deterioration
Patient	Parents or guardian	Substance use	Patient portal	Avoiding care, overdose, or continued addiction
Patient	Parents or guardian	Sexual history or reproductive health	Patient portal	Avoiding care, complications from sexually transmitted infection, or infertility
Patient	Parents or guardian	Gender management or identity	Patient portal	Avoiding care, delay in gender reassignment, or psychological impact
Patient	Parents, guardian, or abuser	Violence or abuse (physical or sexual)	Patient portal	Avoiding care, continued abuse, complications, or death
Patient	Parents or guardian	Complex social situations	Patient portal	Avoiding care or delayed care
Patient	Parents or guardian	Neglect	Patient portal	Avoiding care, delayed care. or continued neglect
Child/Adolescent	Parents or guardian	Foster or custody issues	Patient portal	Avoiding care, delayed care, or family strife
Child/Adolescent	Parents or guardian	Misattributed paternity	Patient portal	Avoiding care, delayed care, or family strife
Parent/Care Giver / Legal Guardian	Patient, other parent, or other care giver	Perinatally acquired sexually transmitted infection	Patient portal	Avoiding care, delayed care, or family strife
Parent/Care Giver / Legal Guardian	Patient, other parent, or other care giver	Substance abuse	Patient portal	Avoiding care, delayed care, or family strife
Parent/Care Giver / Legal Guardian	Patient, other parent, or other care giver	Parent or caregiver's mental health	Patient portal	Avoiding care, delayed care, or family strife
Parent/Care Giver / Legal Guardian	Patient, other parent, or other care giver	Violence, abuse, or legal problems	Patient portal	Avoiding care, delayed care, or family strife
Parent/Care Giver / Legal Guardian	Patient, other parent, or other care giver	Misattributed paternity	Patient portal	Avoiding care, delayed care, or family strife
Parent/Care Giver / Legal Guardian	Patient, other parent, or other care giver	Stress associated with chronic care	Patient portal	Family strife or mistrust
Provider	Patient, other parent, or other care giver	Patient or family disagreement with provider	Patient portal	Delayed or missing documentation
Provider	Patient, other parent, or other care giver	Neglect or abuse	Patient portal	Lawsuit or unsafe environment for the provider

Inadvertent Disclosure of Medically Relevant Information Obtained From Proxies

There may be situations in which health care providers may document pertinent information that they receive from parents, relatives, and legal representatives, which may adversely affect the patient, parent, or legal representative or damage relationships when disclosed (Table 1). For example, parents may disclose their difficulty in coping with an adolescent's chronic illness to a provider who documents it in the adolescent's chart. This information could then be seen by the adolescent in their patient portal and affect the parent–child relationship. Another example is if a parent discloses information about a drug use during pregnancy or perinatally acquired STI to the pediatrician caring for the newborn. This information would be accessible through the infant's electronic record by other users, such as the other parent. In both situations, disclosing medically relevant information may be disincentivized for fear of its discovery by another person having access to the medical record.

Health Care Provider–Patient Relationship in Difficult Diagnostic Dilemmas

Disclosure of information can adversely affect health care provider–patient relationships, especially when there is disagreement between the health care provider and the parents or patient (Table 1). In functional disorders where the medical work up does not demonstrate an organic etiology for the complaint, the parents or patient may believe otherwise. For example, when the defined Rome's Criteria of Functional Abdominal Pain fits a patient's symptoms, parents or the patient may disagree with this diagnosis. In similar cases where the relationship among the patient, family, and health care provider is critical to helping the patient improve, documenting this information could further damage a fraught or tenuous relationship with the health care provider. Although providers should hold themselves to high standards for documenting information in the EHR, providers should not feel pressured to augment their documentation for fear of their medical opinion offending patients or proxies. This can be the case when child abuse is in a differential diagnosis, and documentation of this in the child's record may adversely affect the relationship between parent and health care provider if the parent feels unfairly accused or judged. There are situations where abuse is in a differential diagnosis, albeit with a very low index of suspicion, or where a provider may want to document that they have thought of but ruled out abuse or neglect. In these cases, it is unlikely the information will be compiled in "reasonable anticipation of, or for use in, a civil, criminal, or administrative action or proceeding" [32], which is clearly protected and eligible for legal blocking by HIPAA, and the remainder of the documented information may be of interest to the patient or proxy. The limited capabilities of data segmenting technology create an awkward or burdensome situation for providers.

Older Adult Patients

Overview

The complexity of care and the large number of comorbidities and treatments associated with aging make the electronic patient portal an attractive tool for persons with multiple health conditions. However, many older adults feel uncomfortable or ill-equipped using technology and rely on their caregivers for their health care–related tasks, necessitating proxy portal access. Less than 20% of US hospitals that allow caregiver proxy access also allow patients to filter or partially block the EHI passed on to their proxies [33]. Therefore, older adults are faced with many of the same challenges and potential harms that adolescents may experience.

The Older Adult Patient–Caregiver Relationship

Despite an increase in STIs among adults over the age of 65 years, many older adults are reluctant to share a recent sexual encounter [34] with their health care provider, knowing that this information will be available to caregivers. Syphilis, which is a treatable condition, can mimic dementia and neurocognitive disorders in late stages of the disease if the diagnosis is missed. Similarly, older patients may withhold health information regarding mental health (including depression) and elder abuse (physical, sexual, emotional, neglect, abandonment, financial, and self-neglect) from their health care providers for fear of their proxy finding out (Table 1). Again, the emphasis on broad access may paradoxically erode the patient's control over who can access their data. One potential solution may be allowing patients to block all information related to a specific topic from all users of the patient portal, including themselves, and unblock it again when they become sole users of the portal.

Inadvertent Disclosure of Medically Relevant Information Obtained From Caregivers

Caregivers may disclose emotional, physical, or mental exhaustion leading to burnout. If this information is documented and shared with the older adult patient, unintended consequences are feelings of guilt, overburdening, and depression (Table 1).

Of note, there is a clause in the HIPAA Privacy Rule that specifically addresses keeping third-party information confidential. According to this clause [35]:

> *Any information disclosed to the provider by another person who is not a healthcare provider that was given under a promise of confidentiality (such as that shared by a concerned family member), may be withheld from the patient if the disclosure would be reasonably likely to reveal the source of the information.*

Since the Cures Act defers to the HIPAA, this clause should be applicable under the Cures Act; however, this is likely not well-known or understood across institutions.

In some cultures, it is common practice for caregivers to withhold negative information such as the diagnosis of a cancer or a terminal illness. Caregivers are also frequently surrogate decision-makers and may for many reasons ask a health care provider to withhold a diagnosis [36]. Although we consider

disclosure to the patient as the ethically preferable choice, we acknowledge that the inability to block information may not align with the cultural norms of certain patient groups [37]. The patient may also desire information blocking, such as when an older patient is afraid that disclosure of a new diagnosis of cancer or recurrence may burden their caregiver or lead to caregiver burnout.

Health Care Provider–Patient Relationship in Difficult Diagnostic Dilemmas

Maintaining a good relationship with patients is critical for health care providers taking care of older adults, as dynamic shifts in health often require changing or transitioning goals of care. Even neutral personal descriptors such as "elderly" in a note can make patients feel judged and perceived themes of disrespect, errors, and surprises can lead to straining of a patient–provider relationship [38]. For example, the term palliative is often misinterpreted for end-of-life care when in fact the goal is symptom and quality of life improvement for any serious illness (even curative ones), irrespective of prognosis. Further, the National Center for Educational Statistics reports that 21% of adults in the United States (~43 million) are illiterate or functionally illiterate [39]. Misinterpreting documentation may prevent older adults from seeking care to relieve symptoms and stress and align treatment options with their goals. One unintended, but positive, consequence of the information blocking rules might be that it encourages providers to be more vigilant in their documentation to achieve language that is both medically accurate and affirming of the patient's dignity.

Where Information Blocking Went Too Far

Although the potential adverse outcomes previously discussed do not meet the Cures Act Final Rule definition of harm, in some cases, releasing this information may violate the foundational principles of a trusting provider–patient relationship.

Information Blocking in the Multiuser EHR

While the Cures Act final has made it easier to access information electronically, this increase in access is not accompanied with the requisite technical advances to block access to data in appropriate circumstances. In the situations when information blocking can be legally used, strategies are limited in number and capability, especially in the context of a multiuser or proxy access. Information blocking is technically and logistically challenging, and the burden is placed upon the documenting health care providers to determine what EHI is and is not appropriate to block and to whom. For some health care providers, such as those practicing in adolescent medicine, family medicine, general pediatrics, pediatric subspecialties, internal medicine, and geriatrics, navigating information blocking may be a routine experience depending on patient needs.

At the institutional level, the hospital system can deactivate proxy access; however, this may be burdensome and can be delayed depending upon institutional implementation (eg, a health care provider clicking a button in the EHR versus contacting health information management and placing a ticket for a request to be completed). The ONC exceptions emphasize that information blocking should be no broader than necessary. It will be an infrequent occurrence that a patient or proxy is completely blocked from accessing all EHI and more common that the blocking will occur on a data-element-by-data-element (clinical documentation, laboratory tests, imaging, etc) basis. This may create a substantial burden for the health care institutions and be prone to user errors. Additionally, the absence of information may be conspicuous when a patient or proxy who usually receives information does not. There is an evolving standard called "Data Segmentation for Privacy" (DS4P) where a health care provider could mark portions of a note to be blocked from access; however, the adoption of this standard is minimal [40,41].

Beyond institutional policies and EHR technical capabilities, the health care provider can adopt new documentation workflows when information blocking is legally acceptable. For example, the health care provider could create one note that is appropriate to share with all users and another that includes the information which is then blocked (ie, shadow charting); however, this solution is time-consuming and burdensome and unlikely to be adopted as clinical documentation has already been shown to be a significant contributor to burnout among health care providers [42-46]. Further, duplicate documentation would also be error-prone, jeopardizing safety and creating additional work and confusion for other health care providers on the treatment team relying on documentation to support patient care. Health care providers may choose to avoid caring for patients who are more prone to these complicated situations.

Where information blocking is not acceptable, the health care provider, not wanting to damage a relationship or breach confidentiality, may decide to stop documenting certain information. This is a potentially dangerous practice that could affect medical care, reduce accurate billing, and result in incomplete communication about the patient's medical history with other health care providers.

Conclusions and Recommendations

The Cures Act Final Rule is undoubtedly necessary to facilitate significant improvements in patient care and innovation; however, in some cases of a multiuser or proxy access situation, the Information Blocking Provision conflicts with the standard that health care providers hold themselves to in the United States. Additionally, applying these exceptions to the Information Blocking Provision in legally acceptable cases will be burdensome and could lead to increased burnout among health care providers. Paradoxically, providing patients more control over their data may actually jeopardize their control and privacy in some scenarios. Although breaching confidentiality and damaging the patient–provider relationship will not necessarily cause substantial harm as defined by the text of the Final Rule, it may cause unnecessary anguish, limit the quality of care, or cause a patient to forgo or delay care, and lead to increased morbidity. Additionally, the privacy exception may be underutilized as it necessitates patients and providers be

educated on the application of this rule and an institution's policies and procedures. In light of these concerns, we recommend that ONC provide more detailed guidance both on their website and targeted outreach to health care providers caring for patients in the adolescent health setting and other multiuser or proxy access situations. As clause 171.202 (b) of the Cures Act allows institutions to develop policies around information blocking, we encourage ONC to develop and publish sample policies that institutions may use or modify. Such guidance should outline the exact processes by which a patient can opt out of their health data being shared with a proxy user using the privacy exception and detail how providers can best guide patients through decision-making without the fear of being in violation of the information blocking rules. Where data segmentation for privacy is not feasible, we recommend that ONC considers carving out an option for providers to return to traditional sharing options to prevent breaches of privacy. We also urge ONC to interpret the privacy exception broadly and not penalize hospitals or providers for information blocking when proxy access is the reason for the information blocking. We suggest that ONC and researchers carefully monitor charting habits in these multiuser or proxy access situations by studying how often patients use the privacy exception compared with single-user EHR access scenarios, how much time is spent documenting for these scenarios, and how much shadow-charting is taking place. We also suggest researchers carefully monitor the effect of information blocking on patient, provider, and proxy relationships. Additionally, we recommend limited penalties on health care providers in multiuser or proxy access situations during the implementation process of the Cures Act Final Rule until technological capabilities advance to better segment notes and block them from certain users.

Conflicts of Interest

ML has been compensated for presentations for the Texas Pediatric Society, for consulting on modules to educate providers for Texas Health Steps, for time working with the Texas Essential Knowledge and Skills Proclamation 2022 Review Adoption Review Panel, and as a consultant for the Texas Child Psychiatry Access Network. She was also awarded support funds through the Technology and Adolescent Mental Well Being (TAM) youth Advisory Board (funding to start in January 2022), received a grant from Texas Pediatric Society Foundation, and is a Nexplanon trainer for Organon. CUL owns shares in Celanese and Markel. The remaining authors have no conflicts to declare.

References

1. P.L. 114–255, Approved December 13, 2016 (130 Stat. 1033). Social Security. 2016. URL: https://www.ssa.gov/OP_Home/comp2/F114-255.html#:~:text=An%20Act%20to%20accelerate%20the,%E2%80%9C21st%20Century%20Cures%20Act%E2%80%9D [accessed 2022-02-09]
2. 21st Century Cures Act: Interoperability, Information Blocking, and the ONC Health IT Certification Final Rule: Summary of Information Blocking and Eight Exceptions. CHiME Public Policy. URL: https://chimecentral.org/wp-content/uploads/2020/06/061420_CHIME-Information-Blocking-Cheat-Sheet-FINAL-RULE1-1.pdf [accessed 2022-02-09]
3. About ONC. The Office of the National Coordinator for Health Information Technology (ONC). URL: https://www.healthit.gov/topic/about-onc [accessed 2021-03-20]
4. About ONC's Cures Act Final Rule. The Office of the National Coordinator for Health Information Technology (ONC). URL: https://www.healthit.gov/curesrule/overview/about-oncs-cures-act-final-rule [accessed 2021-03-20]
5. What ONC's Cures Act Final Rule Means for Patients. The Office of the National Coordinator for Health Information Technology. URL: https://www.healthit.gov/curesrule/what-it-means-for-me/patients [accessed 2021-03-20]
6. What ONC's Cures Act Final Rule Means for Clinicians and Hospitals. The Office of the National Coordinator for Health Information Technology. URL: https://www.healthit.gov/curesrule/what-it-means-for-me/clinicians [accessed 2021-03-20]
7. What ONC's Cures Act Final Rule Means for Health IT Developers. The Office of the National Coordinator for Health Information Technology. URL: https://www.healthit.gov/curesrule/what-it-means-for-me/health-it-developers [accessed 2021-03-20]
8. Michael Lipinski, Mark Knee. 21st Century Cures Act: Interoperability, Information Blocking, and the ONC Health IT Certification Program Proposed Rule. US Department of Health and Human Services. URL: https://www.healthit.gov/sites/default/files/facas/2019-03-22_ONC_Cures_Act_NPRM_IB_%20HITAC_IB_Task_Force_508.pdf [accessed 2022-02-12]
9. Everyone on the Same Page. OpenNotes. URL: https://www.opennotes.org/ [accessed 2021-03-20]
10. Hägglund M, DesRoches C, Petersen C, Scandurra I. Patients' access to health records. BMJ 2019 Oct 02;367:l5725. [doi: 10.1136/bmj.l5725] [Medline: 31578180]
11. United States Core Data for Interoperability (USCDI). The Office of the National Coordinator for Health Information Technology (ONC). URL: https://www.healthit.gov/isa/united-states-core-data-interoperability-uscdi [accessed 2021-03-20]
12. Information Blocking FAQs. The Office of the National Coordinator for Health Information Technology. URL: https://www.healthit.gov/curesrule/resources/information-blocking-faqs [accessed 2021-03-20]
13. New Applicability Dates included in ONC Interim Final Rule. The Office of the National Coordinator for Health Information Technology. URL: https://www.healthit.gov/cures/sites/default/files/cures/2020-10/Highlighted_Regulatory_Dates_Information_Blocking.pdf [accessed 2021-03-20]

14. Certified Health IT. The Office of the National Coordinator for Health Information Technology. URL: https://www.healthit.gov/playbook/certified-health-it/ [accessed 2021-03-20]
15. Information Blocking. The Office of the National Coordinator for Health Information Technology. URL: https://www.healthit.gov/curesrule/final-rule-policy/information-blocking [accessed 2021-03-20]
16. Cure Act Final Rule: Information Blocking Exceptions. The Office of the National Coordinator for Health Information Technology. URL: https://www.healthit.gov/sites/default/files/cures/2020-03/InformationBlockingExceptions.pdf [accessed 2021-03-12]
17. Dufendach KR, Eichenberger JA, McPheeters ML, Temple MW, Bhatia HL, Alrifai MW, et al. Core Functionality in Pediatric Electronic Health Records. Rockville, MD: Agency for Healthcare Research and Quality; 2015.
18. Council on Clinical Information Technology. Policy Statement--Using personal health records to improve the quality of health care for children. Pediatrics 2009 Jul;124(1):403-409. [doi: 10.1542/peds.2009-1005] [Medline: 19564327]
19. Society for Adolescent Health and Medicine, Gray SH, Pasternak RH, Gooding HC, Woodward K, Hawkins K, et al. Recommendations for electronic health record use for delivery of adolescent health care. J Adolesc Health 2014;54(4):487-490. [doi: 10.1016/j.jadohealth.2014.01.011] [Medline: 24656534]
20. Schapiro NA, Mihaly LK. The 21st Century Cures Act and Challenges to Adolescent Confidentiality. J Pediatr Health Care 2021;35(4):439-442. [doi: 10.1016/j.pedhc.2021.03.005] [Medline: 33865680]
21. 21st Century Cures Act: Interoperability, Information Blocking, and the ONC Health IT Certification Program.: Federal Register URL: https://www.federalregister.gov/documents/2020/05/01/2020-07419/21st-century-cures-act-interoperability-information-blocking-and-the-onc-health-it-certification [accessed 2022-02-12]
22. Delbanco T, Walker J, Bell SK, Darer JD, Elmore JG, Farag N, et al. Inviting patients to read their doctors' notes: a quasi-experimental study and a look ahead. Ann Intern Med 2012 Oct 02;157(7):461-470 [FREE Full text] [doi: 10.7326/0003-4819-157-7-201210020-00002] [Medline: 23027317]
23. Esch T, Mejilla R, Anselmo M, Podtschaske B, Delbanco T, Walker J. Engaging patients through open notes: an evaluation using mixed methods. BMJ Open 2016 Jan 29;6(1):e010034 [FREE Full text] [doi: 10.1136/bmjopen-2015-010034] [Medline: 26826154]
24. Gerard M, Chimowitz H, Fossa A, Bourgeois F, Fernandez L, Bell SK. The Importance of Visit Notes on Patient Portals for Engaging Less Educated or Nonwhite Patients: Survey Study. J Med Internet Res 2018 May 24;20(5):e191 [FREE Full text] [doi: 10.2196/jmir.9196] [Medline: 29793900]
25. Bialostozky M, Huang JS, Kuelbs CL. Are You In or Are You Out? Provider Note Sharing in Pediatrics. Appl Clin Inform 2020 Jan;11(1):166-171 [FREE Full text] [doi: 10.1055/s-0040-1701679] [Medline: 32131116]
26. Sinha S, Puttagunta R, Vodzak J. Interoperability and Information-Blocking Rules: Implications for Pediatric and Adolescent Health Care Professionals. JAMA Pediatr 2021 Oct 01;175(10):997-998. [doi: 10.1001/jamapediatrics.2021.2077] [Medline: 34309649]
27. Pageler NM, Webber EC, Lund DP. Implications of the 21st Century Cures Act in Pediatrics. Pediatrics 2021 Mar;147(3):e2020034199. [doi: 10.1542/peds.2020-034199] [Medline: 33293349]
28. Ip W, Yang S, Parker J, Powell A, Xie J, Morse K, et al. Assessment of Prevalence of Adolescent Patient Portal Account Access by Guardians. JAMA Netw Open 2021 Sep 01;4(9):e2124733 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.24733] [Medline: 34529064]
29. Parsons CR, Hron JD, Bourgeois FC. Preserving privacy for pediatric patients and families: use of confidential note types in pediatric ambulatory care. J Am Med Inform Assoc 2020 Nov 01;27(11):1705-1710 [FREE Full text] [doi: 10.1093/jamia/ocaa202] [Medline: 32989446]
30. Carlson J, Goldstein R, Hoover K, Tyson N. NASPAG/SAHM Statement: The 21st Century Cures Act and Adolescent Confidentiality. J Adolesc Health 2021 Feb;68(2):426-428. [doi: 10.1016/j.jadohealth.2020.10.020] [Medline: 33541602]
31. Reddy DM, Fleming R, Swain C. Effect of mandatory parental notification on adolescent girls' use of sexual health care services. JAMA 2002 Aug 14;288(6):710-714. [doi: 10.1001/jama.288.6.710] [Medline: 12169074]
32. Under what circumstances may a covered entity deny an individual's request for access to the individual's PHI? U.S. Department of Health & Human Services. URL: https://www.hhs.gov/hipaa/for-professionals/faq/2046/under-what-circumstances-may-a-covered-entity/index.html [accessed 2021-12-05]
33. Latulipe C, Mazumder SF, Wilson RKW, Talton JW, Bertoni AG, Quandt SA, et al. Security and Privacy Risks Associated With Adult Patient Portal Accounts in US Hospitals. JAMA Intern Med 2020 Jun 01;180(6):845-849 [FREE Full text] [doi: 10.1001/jamainternmed.2020.0515] [Medline: 32364562]
34. Bowen V, Braxton J, Davis D. Sexually Transmitted Disease Surveillance 2018. Centers for Disease Control and Prevention. 2018. URL: https://www.cdc.gov/std/stats18/STDSurveillance2018-full-report.pdf [accessed 2022-02-09]
35. Individuals' Right under HIPAA to Access their Health Information 45 CFR § 164.524. U.S. Department of Health & Human Services. URL: https://www.hhs.gov/hipaa/for-professionals/privacy/guidance/access/index.html [accessed 2021-12-04]
36. McCabe MS, Wood WA, Goldberg RM. When the family requests withholding the diagnosis: who owns the truth? J Oncol Pract 2010 Mar;6(2):94-96 [FREE Full text] [doi: 10.1200/JOP.091086] [Medline: 20592784]

37. Ni YH, Alræk T. What circumstances lead to non-disclosure of cancer-related information in China? A qualitative study. Support Care Cancer 2017 Mar;25(3):811-816 [FREE Full text] [doi: 10.1007/s00520-016-3464-z] [Medline: 27832368]
38. Fernández L, Fossa A, Dong Z, Delbanco T, Elmore J, Fitzgerald P, et al. Words Matter: What Do Patients Find Judgmental or Offensive in Outpatient Notes? J Gen Intern Med 2021 Sep;36(9):2571-2578. [doi: 10.1007/s11606-020-06432-7] [Medline: 33528782]
39. Mamedova S, Pawlowski E. Adult Literacy in the United States. National Center for Education Statistics. 2019. URL: https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2019179 [accessed 2022-02-09]
40. McCarthy MW, de Asua DR, Gabbay E, Fins JJ. Off the Charts: Medical documentation and selective redaction in the age of transparency. Perspect Biol Med 2018;61(1):118-129. [doi: 10.1353/pbm.2018.0031] [Medline: 29805152]
41. Security Tags for Sensitive Information. The Office of the National Coordinator for Health Information Technology. URL: https://www.healthit.gov/isa/data-segmentation-sensitive-information [accessed 2021-03-20]
42. Downing NL, Bates DW, Longhurst CA. Physician Burnout in the Electronic Health Record Era: Are We Ignoring the Real Cause? Ann Intern Med 2018 Jul 03;169(1):50-51. [doi: 10.7326/M18-0139] [Medline: 29801050]
43. Shanafelt TD, Dyrbye LN, West CP. Addressing Physician Burnout: The Way Forward. JAMA 2017 Mar 07;317(9):901-902. [doi: 10.1001/jama.2017.0076] [Medline: 28196201]
44. Marmor RA, Clay B, Millen M, Savides TJ, Longhurst CA. The Impact of Physician EHR Usage on Patient Satisfaction. Appl Clin Inform 2018 Jan;9(1):11-14 [FREE Full text] [doi: 10.1055/s-0037-1620263] [Medline: 29298451]
45. West CP, Dyrbye LN, Shanafelt TD. Physician burnout: contributors, consequences and solutions. J Intern Med 2018 Jun;283(6):516-529 [FREE Full text] [doi: 10.1111/joim.12752] [Medline: 29505159]
46. Frintner MP, Kaelber DC, Kirkendall ES, Lourie EM, Somberg CA, Lehmann CU. The Effect of Electronic Health Record Burden on Pediatricians' Work-Life Balance and Career Satisfaction. Appl Clin Inform 2021 May;12(3):697-707. [doi: 10.1055/s-0041-1732402] [Medline: 34341980]

Abbreviations

DS4P: Data Segmentation for Privacy
EHI: electronic health information
EHR: electronic health record
HHS: US Department of Health and Human Services
HIPAA: Health Insurance Portability and Accountability Act
IT: information technology
ONC: Office of The National Coordinator for Health Information Technology
STI: sexually transmitted information
TAM: Technology and Adolescent Mental Well Being

Edited by T Leung; submitted 06.10.21; peer-reviewed by N Pageler, J Hron; comments to author 06.11.21; revised version received 07.12.21; accepted 26.12.21; published 17.02.22.

Please cite as:
Arvisais-Anhalt S, Lau M, Lehmann CU, Holmgren AJ, Medford RJ, Ramirez CM, Chen CN
The 21st Century Cures Act and Multiuser Electronic Health Record Access: Potential Pitfalls of Information Release
J Med Internet Res 2022;24(2):e34085
URL: https://www.jmir.org/2022/2/e34085
doi:10.2196/34085
PMID:35175207

Viewpoint

Big Data Health Care Innovations: Performance Dashboarding as a Process of Collective Sensemaking

Hilco J van Elten[1], PhD; Sandra Sülz[1], PhD; Erik M van Raaij[1,2], PhD; Rik Wehrens[1], PhD

[1]Erasmus School of Health Policy & Management, Erasmus University, Rotterdam, Netherlands

[2]Rotterdam School of Management, Erasmus University, Rotterdam, Netherlands

Corresponding Author:
Hilco J van Elten, PhD
Erasmus School of Health Policy & Management
Erasmus University
Burgemeester Oudlaan 50
Rotterdam, 3062 PA
Netherlands
Phone: 31 10 408 8869
Email: vanelten@eshpm.eur.nl

Abstract

Big data is poised to revolutionize health care, and performance dashboards can be an important tool to manage big data innovations. Dashboards show the progress being made and provide critical management information about effectiveness and efficiency. However, performance dashboards are more than just a clear and straightforward representation of performance in the health care context. Instead, the development and maintenance of informative dashboards can be more productively viewed as an interactive and iterative process involving all stakeholders. We refer to this process as *dashboarding* and reflect on our learnings within a large European Union–funded project. Within this project, multiple big data applications in health care are being developed, piloted, and scaled up. In this paper, we discuss the ways in which we cope with the inherent sensitivities and tensions surrounding dashboarding in such a dynamic environment.

International Registered Report Identifier (IRRID): RR2-10.2196/16779

(J Med Internet Res 2022;24(2):e30201) doi:10.2196/30201

KEYWORDS

dashboarding; big data; balanced scorecard; performance measurement; key performance indicators; digital health; dashboards; knowledge translation; health information; stakeholders; health care

Introduction

Big data innovations in health care are on the rise, with benefits that include providing clinical intelligence about a patient's risk of future adverse health events [1], empowering patients through big data–driven eHealth applications [2], increasing interest in clinical prediction tools [3], and reducing costs by leveraging big data to detect fraud, abuse, waste, and errors in health insurance claims [4]. Recently, China has used big data technology in its attempts to prevent and control the spread of COVID-19 [5]. All in all, big data innovations in health care promise to address the rising demand for high-quality health care services and to reduce the accompanying costs of care. This makes big data appealing for governments and funding agencies, who are increasingly investing in large-scale, interdisciplinary projects and consortia that seek to develop big data innovations, implement them in the health care field, and ideally demonstrate their desired impact. Such projects usually involve a multitude of stakeholders who, in addition to the project goal, have their own objectives, such as creating social or economic impact, publishing research, or commercializing the innovation. These large investments and diverse goals and objectives call for close monitoring of the overall progress and performance of projects. Such performance monitoring is frequently done utilizing dashboards containing key performance indicators (KPIs).

Dashboards provide a visual overview of the information needed for performance management, and by doing so, facilitate decision-making and project management. In the health care field, the use of dashboards is manifold. Types of dashboard include unit-specific nursing performance dashboards [6], hospital-wide or disease-specific quality and safety diagnostics dashboards [7,8], and population-level maternal and newborn dashboards [9]. Management literature has emphasized the use

of performance dashboards to provide an overview of relevant KPIs and to enable performance management in hospitals [7,10,11]. As such, dashboards are often preferred as managerial instruments to gain insights about progress and performance.

Despite their appeal and aura of rationality, performance dashboards are not simply a clear and straightforward representation of health care performance. Indeed, this point has gained traction within recent accounting literature [12]. Building on the insights from accounting literature, critical management studies, and science and technology studies, we argue that one cannot simply design a dashboard to capture the productivity or efficiency of something as inherently complex and multifaceted as the performance of big data innovations in health care [12]. At the same time, dashboards continue to have enormous appeal, both for funders of research and implementation projects (who perceive them as providing tangible outcomes to account for investments made) and for health care managers (who view them as instruments that can provide valuable information). Therefore, such tools cannot simply be dismissed because of the inherent challenges mentioned above.

In this paper, we introduce our perspective on the development of systems for monitoring big data innovation projects in health care through performance dashboards and discuss caveats concerning these systems. Our perspective focuses on the social dimension of this interactive process, that is, we do not delve into the technical dimensions, such as techniques for data analytics and visualization [13]. Importantly, we view the performance dashboards as the *means* to monitor big data innovation projects. In our paper, they are not the *results* of big data innovations.

We argue that the development of such dashboards does not match the prevalent rationalist narrative usually assumed. Far from being a series of rational decisions, this development can be better understood as a process of collective sensemaking. While the idea of sensemaking has strong roots in organizational psychology, in which the notion is used to study the microfoundations of organizing, we use the notion here to point toward a more general focus on interpretive forms of research that address how meanings are negotiated and constructed [14]. Nevertheless, although we do not build on the full theoretical framework associated with sensemaking, our use also shares affinities with organizational research that describes sensemaking as being "not about truth and getting it right," but instead "about continued redrafting of an emerging story so that it becomes more comprehensive, incorporates more of the observed data, and is more resilient in the face of criticism" [15]. Therefore, we suggest a more pragmatic approach toward the development and use of performance dashboards. In this process, for which we propose the term *dashboarding*, many decisions need to be made jointly with the most important stakeholders. Anything but a linear sequence of decisions, this is a necessarily iterative, recursive process of moving back and forth to find out which indicators are feasible, acceptable, measurable, and informative.

Our viewpoint is based on both our familiarity with the literature regarding the goals, drivers, and problematic aspects of using dashboards as instruments for performance measurement, and on our experiences as researchers and dashboard developers in a large-scale, 3-year European project aimed at developing big data applications in health care, titled "Big Data for Medical Analytics," shortened to "BigMedilytics" (this project was approved by the ethics board of the Erasmus Medical Centre [MEC-2018-056] and the ethics review board of Erasmus University [EA18-01]). Drawing on our experiences in the development, tailoring, and further modification of performance dashboards for 12 pilot projects to develop big data innovations in health care in 8 European countries, our aim is to increase awareness of tensions and to develop sensitivities among academics, clinicians, and practitioners involved in performance measurement of health care innovations, especially those involving big data. We draw on the numerous discussions we have had with the research team and pilot partners, our experiences in organizing various pilot project–specific workshops discussing dashboard designs at various stages, and the many emails, phone calls, and various other interactions we have had with pilot project teams over the last years.

The BigMedilytics Project

Our empirical setting is the 3-year BigMedilytics project, funded by the European Union. This project aims to "enhance patient outcomes and increase productivity in the health sector by applying big data technologies to complex datasets while ensuring security and privacy of personal data" [16]. The entire consortium consists of 35 different entities, ranging from health care providers, technology companies, and insurers to research institutes and universities. At the core of the consortium are 12 pilot projects that develop big data innovations in health care in 8 European countries, divided into 3 areas: population health and chronic disease management, oncology, and health care services industrialization, as described in the underlying protocol paper [13]. The big data technologies include systems to derive predictive models, clinical decision support systems, and real-time asset location systems.

The Inherent Complexities of Dashboarding

While performance dashboards can provide useful insights into the effectiveness and quality improvement potential of big data innovations, rational, straightforward design and implementation of these dashboards is problematic for several reasons. Building on literature in accounting, critical management studies and science and technology studies, we identify 4 important insights that deeply problematize the idea that dashboards simply "capture" performance. First, performance is an inherently debated and complex concept [17,18]. At the core of health care systems are health care organizations that seek to fulfill multiple, sometimes even conflicting aims. As such, tensions are likely to arise between what counts as "performance" in different contexts. This makes performance monitoring and performance management inherently complex. Second, there is always a trade-off between validity and feasibility. The search for exhaustive validation of a dashboard comes at the cost of practical feasibility; when measurement aims to provide a flawless map of the organization's landscape, this often hinders practical utility and enhances strategic behavior [19,20]. Third,

there is some tension between using performance dashboards internally for learning purposes and using dashboards externally as bases for monitoring and accountability [21]. In order to learn from performance dashboards, it is vital to stimulate discussions based on the performance scores. Using performance measurement as an accountability tool is an "external" use of the dashboard, in which those who monitor the project use the dashboard to steer and direct. Using performance measurement as an accountability and control tool is likely to corrode and corrupt the performance indicators, undermining the conditions required for quality improvement [21]. Fourth, performance monitoring is never simply a representation of reality, but also shapes and structures the reality to be acted upon in a process of coconstitution [22,23]. Dashboards never only represent performance or particular aspects of performance; they are also social facts that generate actions and reactions, for example, by defining managerial priorities and by reconfiguring work routines and relationships between actors [24,25]. They are thus "performative" [24,25] in the sense that they do not only represent aspects of organizational performance but also help shape and define the very aspects of performance that come to matter—an aspect that is also recognized within the accounting literature [26].

Three Persistent Tensions

Based on our experiences in the BigMedilytics project, we will describe 3 persistent tensions that we and other actors in the consortium experienced while working toward performance dashboards that met the dual requirements of being useful to pilot partners while allowing for some form of assessment regarding goals and achievements. After this description, we will outline the main implications, reiterate our suggestion that a more pragmatic perspective on dashboarding should be developed, and discuss the implications of this perspective for academics, clinicians, and practitioners involved in the performance measurement of health care innovations.

Tension 1: Navigating Between Divergent Stakeholder Views and Expectations

The consortium that we were a part of included the developers of a range of big data innovation projects, which involved, like many similar projects, various stakeholder groups from academia, government, and industry. At the core of the consortium were 12 big data pilot projects to develop and implement big data innovations in 8 European countries: Austria, France, Germany, Ireland, the Netherlands, Spain, Sweden, and the United Kingdom. These pilot projects were divided into 3 themes: population health and chronic disease management (n=5), oncology (n=3), and health care services industrialization (n=4). Each of the 12 pilot project teams included several members, introducing a multitude of experiences and expertise.

Consortia necessarily bring together a range of stakeholders, who each have their own specific ideas about the goals of the project. Many innovation projects, not just those organized as consortia, involve internal or external funding. In order to apply for funding, stakeholders need to align goals with each other. An inherent feature of the application and alignment process is time constraint, which hampers the ability to acquire all the information needed in order to make a fully informed decision. Consequently, one must settle on a goal that is deemed acceptable for all groups given the information available at the time. This necessitates broadly defined goals that require further specification over time. In our case, the BigMedilytics consortium aligned on the broad and ambitious aim to improve health care in Europe through big data and "demonstrate an increase in healthcare productivity between 20% and 63%" [27].

But as time unfolded, it became evident that people attributed different meanings to the term *productivity*, which affected the way they conceptualized and operationalized it. One prominent definition of productivity considers it to be the ratio of input to output and to thereby express how efficiently resources are used [28,29]. Defining the overall project goal in such strong economic and engineering language, however, did not match the perception of how individual pilot projects could contribute to achieving productivity gains by deploying big data. With the big data innovations organized in 12 very different pilot projects, various idiosyncratic features needed to somehow be reflected in how the productivity of the projects was measured. A one-size-fits-all approach was considered inappropriate and a certain degree of divergence was deemed inevitable. One might raise the argument that this divergence was a consequence of the heterogeneity of the pilot projects, and consider that homogeneous pilot projects would not face this tension. While this argument has some appeal, we know from the literature that even in similar settings, different opinions prevail on what constitutes productivity [30], performance [31], and quality of care [32]. Consequently, some divergence is inevitable if one is to maintain stakeholders' support and engagement. Stakeholders need to see their pilot project–specific contribution to the overall goal adequately reflected.

This divergence needs to be managed. The more pilot project–specific idiosyncrasies are considered, the better pilot project–specific developments can be monitored and the more precisely we can formulate a pilot project–specific conclusion. By the same token, the more pilot project–specific idiosyncrasies are considered, and the more pilot projects diverge, the more challenging it becomes to converge again and draw conclusions across pilot projects. This indicates a tension: if divergence amongst pilot projects is too high, goal congruence might be compromised. On the other hand, not allowing any divergence would amount to the one-size-fits-all approach that was already considered inappropriate. Managing this tension thus calls for a compromise that allows for some divergence but at the same time enables alignment between stakeholders and convergence toward the overall project goal.

Striving for convergence but simultaneously allowing for necessary divergence resulted in the following approach: we applied a broader angle and considered performance along multiple dimensions. While the multi-dimensionality of the performance framework thus increased the likelihood that stakeholders considered the framework acceptable, it initiated a discussion on what type of performance dimension adequately captured the objectives and intentions of the big data innovations. This called for a compromise: we decided to align on dimensions broad enough that they applied to multiple pilot projects (in step 1), but also recognized that we needed to give pilot projects the possibility to define KPIs within these

dimensions that adequately captured their unique qualities (in step 2). We used a balanced scorecard (BSC) [33] to monitor multiple performance dimensions and tailored its design by determining pilot project–specific KPIs. The BSC [33,34] is a widely applied framework, primarily in for-profit business organizations, and consists of a set of KPIs that gives top managers a fast but comprehensive view of the organization. The set of KPIs includes financial performance measures that show the results of actions already taken. The BSC complements the financial measures with operational, process, and quality measures of various performance dimensions that are the drivers of future financial performance.

Although popularized in for-profit business organizations, the application and study of BSCs in the health care industry is not new. In the health care industry, most prominent descriptions of BSCs refer to monitoring quality aspects of performance. For example, Chong et al [35] found that "measuring the quality of a hospital is an important but exceedingly difficult task. Different methods of capturing quality have been devised, including composite scores of compliance with various quality indicators to the adoption of BSC techniques from the business world." Also, Fernando et al [36] support this idea by claiming that scorecards are used by institutions "for the purposes of monitoring clinical performance and driving quality improvement."

Tension 2: Navigating Between Timely and Meaningful Data Collection

For each of the performance dimensions of the BSC (ie, patient satisfaction, patient outcomes, process outcomes, and financial outcomes), we determined pilot project–specific KPIs in close collaboration with the pilot project stakeholders via an iterative procedure. For example, pilot projects that related to population health, chronic disease, and oncology often included long-term financial measures (eg, the projected cost of care over 10 years) next to short-term measures, such as the average cost per patient. In most pilot projects, patients were expected to be directly impacted by the big data innovation. In these cases, patient satisfaction surveys and patient-reported measures were relevant to measure the dimension "patient satisfaction." Mortality was a typical measure for the dimension "patient outcomes." In contrast, the pilot projects related to industrialization did not measure KPIs for the "patient satisfaction" dimension, as patients were not directly impacted by the big data innovations, but instead focused on process outcomes (such as completion time for diagnoses).

Through this collaboration we aimed to ensure the relevance of the selected KPIs to the pilot projects, as well as increase commitment of the pilot project teams to the dashboard. This commitment was key as the pilot project teams were supposed to report on these KPIs every 6 months. The suggestion from the project consortium was to rely on standardized lists of KPIs. These KPIs were often already available in the reporting and management systems used in the pilot projects or the organizational bodies (eg, hospitals) to which the pilot projects belonged. Such availability would make reliable data collection very efficient. However, efficient data collection of readily available KPIs is not the same thing as valid and meaningful data collection of KPIs that capture the actual performance and progress of pilot projects.

Therefore, we engaged in a discussion with the pilot project teams about the strategic aims of the pilot projects to make an initial assessment of priorities and feasibility. First, we organized a series of workshops where we presented the 4 dimensions to pilot project stakeholders and let the stakeholders openly brainstorm potential KPIs for each of these dimensions. Then, we compared the KPIs across pilot projects to identify similarities and differences and inquired whether KPIs that were suggested in other pilot projects might also be relevant for the pilot project in question. As a final step in the development of the KPI dashboard, the researchers and the pilot project stakeholders agreed on a set of pilot project–specific KPIs that were deemed relevant to the pilot project in question and for which reliable data were expected to be periodically collected.

After a baseline measurement of the KPIs, which provided a crucial anchor point for benchmarking the pilot project's subsequent performance, the KPIs were supposed to be updated periodically using the same operating procedures and definitions. Receiving periodic updates has, however, frequently been challenging. Reasons for this were manifold: updating the KPIs was not deemed meaningful at that point in time; pilot projects were partially dependent on other entities to deliver the data, which prevented timely data submission; databases were not updated as frequently as intended; information technology systems were changing over time, which required modifying the operating procedures to gather the KPI data; and some KPIs turned out to be unreliable and needed to be revised whereas other KPIs—despite all good intentions—simply could not yet be measured due to a lack of data. Consequently, the set of KPIs changed over time, altering the design of the dashboard and reducing the reliability of the measurement over time.

In contrast to the idea that KPIs can capture performance and progress over time, this project made clear that a set of KPIs—even if tailored to a specific context—cannot be considered to be static, but must rather be thought of as dynamic, evolving, and changing. This holds true specifically for big data innovations in which more and more learnings are generated as time passes and new data become available. The trade-off between validity, reliability, and feasibility becomes apparent: KPIs that have initially been considered a valid indicator of performance might be more reliable due to the amount of data available, but they may also no longer be considered to be valid because they fail to capture state-of-the-art performance. Again, this tension points toward the need to understand KPI dashboards in process terms: as a dynamic process of "dashboarding" that requires adaptations over time in order to remain informative given the state of the art.

Tension 3: Navigating Between Different Dashboarding Needs and Purposes

Through collaboration with each individual pilot project team, we wanted to develop dashboards that would do justice to the idiosyncratic nature of the pilot projects and would be loaded with relevant and timely datapoints for an informative performance analysis. Performance analyses can be informative for the members of each pilot project team, enabling them to

learn and improve, and for those monitoring progress across pilot projects. Serving both these needs, however, turned out to be difficult. Monitoring progress across pilot projects required comparability that could only be achieved by choosing a set of KPIs that was a compromise; they did not fit individual projects perfectly, making them a suboptimal choice. Learning and improving within a pilot project, however, requires selecting KPIs that are as closely related to the innovation as possible. Since big data innovations themselves do not improve health care directly, but rather contribute to improvements by changing the information used in decision-making, the way the innovation is embedded needs to be reflected in the KPIs. For example, the BigMedilytics "asset management" pilot project deployed a track-and-trace interface that professional caregivers used to locate medical equipment. As such, the interface was directly involved in a caregiver's search process and could thereby directly influence the health care process. There were also pilot projects, however, where such direct embedding was not the case. The purpose of the "stroke workflow" pilot project was to use big data to identify bottlenecks in the workflow. As such, big data was used as a diagnostic tool to identify problematic areas in the health care process. This identification step revealed changes in the health care process that needed to be made, yet the big data innovation itself was not directly involved in health care delivery. Similar arguments pertain to other BigMedilytics pilot projects, showing that the way in which the innovation is used and how it affects the health care process differ in specific situations.

With pilot project members closely involved in the KPI selection process, one would expect the dashboards to be used for learning and improvement purposes. However, we did not observe strong indications of this. In our experience over the years, the dashboards were more often perceived as a managerial necessity for which KPI data needed to be provided for monitoring purposes. Ironically, we also did not observe strong indications of dashboards being used to oversee progress across pilot projects, which seems to indicate that the overall purpose of the dashboard remained unclear and that its potential to facilitate learning and monitor progress was insufficiently realized.

Implications

Based on our experience with developing performance dashboards for 12 big data innovation pilots, we have argued for the need to develop a more pragmatic, process-based perspective on performance dashboards. We describe this as the notion of "dashboarding." This perspective recognizes that despite its rationalistic aura, developing, tailoring, and modifying performance dashboards is in essence an unpredictable, messy, and iterative process that (1) involves a wide range of stakeholders with often diverging goals and expectations, (2) calls for situation-specific assessments of the balance between efficient and meaningful data collection, and (3) comes with struggles with hard-to-reconcile demands, such as the need to monitor achievements across pilot projects to account for investments made versus the need to provide tailored insights to help specific pilot project teams evaluate and improve their performance.

What are the implications of this perspective for academics, clinicians, and practitioners involved in the performance measurement of health care innovations? The first implication is that those involved in the process of dashboarding need to develop the political sensitivity to acknowledge and manage differences in interests and objectives among various stakeholders that are directly involved in and indirectly affected by the dashboard. Importantly, this goes beyond the idea of cocreation with stakeholders. While incorporating stakeholders early on is an important condition for generating support, it is by no means a panacea [37]. With an increasing number of stakeholders involved, the diversity in interests and expectations is likely to increase as well. Obviously, this may cause tensions if interests and expectations are diverging or even conflicting. In line with related literature [7,37], we therefore suggest involving stakeholders in co-designing dashboards, but we also stress that this requires careful expectation management, sensitivity toward different needs and requirements, and persistence in navigating between different perspectives and interests. Collaborative design brings together specialists and generalists from various backgrounds who share knowledge of the design process as well as the design content in order to create shared understanding of both aspects [38]. Such principles of collaborative design can be combined with principles of participatory design, which refers to the participation of prospective users, who become true participants and not just informants in the design process [39]. We involved future dashboard users from different disciplines (including data scientists and medical professionals) to some extent but would recommend more extensive and more explicit use of collaborative and participatory design principles in the dashboarding process.

The second implication of the pragmatic perspective we propose relates to the iterative nature of dashboarding. Developing dashboards iteratively through multiple stages of refinement does not align well with the logic of "projectification" that often underlies large-scale programs like the BigMedilytics program [40,41]. The notion of projectification refers to a mode of science governance that sets the concept of the "project" as its basic organizing principle [42]. As such, it highlights not only the ubiquity of the project format, but also points to underlying instrumental reasoning and a rationalistic attitude toward predictability (eg, an emphasis on activities leading to "milestones" and "deliverables") that is often considered to be in tension with other values of research and innovations (such as academic freedom and creativity) [40,43]. This attitude is at odds with the dynamic and continuously evolving environment organizations find themselves in. Far from a controlled, experimental setting, pilot projects exist among a swirl of other projects, initiatives, developments, and changes. In the BigMedilytics program, for instance, all pilot projects were confronted with two major changes: the introduction of the General Data Protection Regulation and the COVID-19 pandemic. The main implication of this is that researchers and dashboard developers would benefit from building in more leeway for interim changes. Thinking through the implications of this to the fullest extent would also require funders and other policy makers to think about innovative ways of funding

research and technology development that do not take the logic of projectification as a pre-eminent reality.

The third implication relates to the need for flexibility in dashboarding. Dashboards are used for different purposes that range from providing external accountability and internal benchmarking to enabling improvement initiatives [7]. Dashboards are also frequently considered to be informational tools that stimulate discussion and critical reflection [6]. As such, dashboards need to be flexible enough to cater to multiple purposes and different stakeholder needs [7]. This flexibility requires a specific view on progress monitoring via performance dashboards. Instead of striving for comparability across dashboards in a volatile environment (which is particularly relevant for big data as well as health care), progress monitoring can also be achieved with a focus on continuous improvement. If each project team develops its own set of KPIs (with some steering, as described above), a dashboard will emerge that will be deemed informative by the project team. Project teams should also regularly reflect on whether the set of KPIs is still informative and whether adaptations are necessary. The reasons for incorporating a new KPI or dropping an old one reveal important insights about the progress that can even go beyond the indications derived from comparing KPI scores across project teams or time.

Future research might investigate which specific approaches, tools, techniques, programming languages, and design environments are most effective for such a collaborative, iterative and flexible dashboarding process. Successful cases of dashboard development, including details on the technologies used, are available in the literature [7,44-46], but there remains a need for a systematic comparison of dashboarding approaches.

Conclusion

Dashboarding is a dynamic process that features various tensions. Instead of neglecting these tensions, we plead for reflection upon them and navigation through them. Our recommendations therefore do not come in the form of a magic bullet and do not offer clear-cut solutions but are rather a description of 3 sensitivities that performance dashboard designers should develop to handle the tensions involved in the process of dashboarding. Capitalizing on these sensitivities will lead to dashboarding that is iterative, integrative, and informative.

Acknowledgments

This paper is part of a series of papers related to the EU Horizon 2020 project “Big Data for Medical Analytics” (abbreviated as “BigMedilytics”). This work was supported by Horizon 2020 grant number 780495.

Disclaimer

Any dissemination of results here presented reflects only the authors’ views. The European Commission is not responsible for any use that may be made of the information it contains.

Authors' Contributions

HJvE and SS wrote the first draft of the paper. EMvR and RW worked on revisions of the paper. All authors have provided feedback on multiple draft versions. HJvE finalized the revisions.

Conflicts of Interest

None declared.

References

1. Lin Y, Chen H, Brown RA, Li S, Yang H. Healthcare Predictive Analytics for Risk Profiling in Chronic Care: A Bayesian Multitask Learning Approach. MISQ 2017 Feb 1;41(2):473-495. [doi: 10.25300/misq/2017/41.2.07]
2. Lee EWJ, Viswanath K. Big Data in Context: Addressing the Twin Perils of Data Absenteeism and Chauvinism in the Context of Health Disparities Research. J Med Internet Res 2020 Jan 07;22(1):e16377 [FREE Full text] [doi: 10.2196/16377] [Medline: 31909724]
3. Barda AJ, Horvat CM, Hochheiser H. A qualitative research framework for the design of user-centered displays of explanations for machine learning model predictions in healthcare. BMC Med Inform Decis Mak 2020 Oct 08;20(1):257 [FREE Full text] [doi: 10.1186/s12911-020-01276-x] [Medline: 33032582]
4. Srinivasan U, Arunasalam B. Leveraging Big Data Analytics to Reduce Healthcare Costs. IT Prof 2013 Nov;15(6):21-28. [doi: 10.1109/mitp.2013.55]
5. Wu J, Wang J, Nicholas S, Maitland E, Fan Q. Application of Big Data Technology for COVID-19 Prevention and Control in China: Lessons and Recommendations. J Med Internet Res 2020 Oct 09;22(10):e21980 [FREE Full text] [doi: 10.2196/21980] [Medline: 33001836]
6. Jeffs L, Beswick S, Lo J, Lai Y, Chhun A, Campbell H. Insights from staff nurses and managers on unit-specific nursing performance dashboards: a qualitative study. BMJ Qual Saf 2014 Dec;23(12):1001-1006. [doi: 10.1136/bmjqs-2013-002595] [Medline: 25193341]

7. Weggelaar-Jansen AMJWM, Broekharst DSE, de Bruijne M. Developing a hospital-wide quality and safety dashboard: a qualitative research study. BMJ Qual Saf 2018 Dec;27(12):1000-1007 [FREE Full text] [doi: 10.1136/bmjqs-2018-007784] [Medline: 29950323]
8. Mane KK, Rubenstein KB, Nassery N, Sharp AL, Shamim EA, Sangha NS, et al. Diagnostic performance dashboards: tracking diagnostic errors using big data. BMJ Qual Saf 2018 Jul;27(7):567-570. [doi: 10.1136/bmjqs-2018-007945] [Medline: 29550767]
9. Dunn S, Sprague AE, Grimshaw JM, Graham ID, Taljaard M, Fell D, et al. A mixed methods evaluation of the maternal-newborn dashboard in Ontario: dashboard attributes, contextual factors, and facilitators and barriers to use: a study protocol. Implement Sci 2016 May 04;11:59 [FREE Full text] [doi: 10.1186/s13012-016-0427-1] [Medline: 27142655]
10. Millar R, Mannion R, Freeman T, Davies HTO. Hospital board oversight of quality and patient safety: a narrative review and synthesis of recent empirical research. Milbank Q 2013 Dec;91(4):738-770 [FREE Full text] [doi: 10.1111/1468-0009.12032] [Medline: 24320168]
11. Ghazisaeidi M, Safdari R, Torabi M, Mirzaee M, Farzi J, Goodini A. Development of Performance Dashboards in Healthcare Sector: Key Practical Issues. Acta Inform Med 2015 Oct;23(5):317-321 [FREE Full text] [doi: 10.5455/aim.2015.23.317-321] [Medline: 26635442]
12. Van Erp W, Roozen F, Vosselman E. The performativity of a management accounting and control system: Exploring the dynamic relational consequences of a design. Scand J Manag 2019 Dec;35(4):101077. [doi: 10.1016/j.scaman.2019.101077]
13. Wehrens R, Sihag V, Sülz S, van Elten HJ, van Raaij EM, de Bont A, et al. Understanding the Uptake of Big Data in Health Care: Protocol for a Multinational Mixed-Methods Study. JMIR Res Protoc 2020 Oct 22;9(10):e16779 [FREE Full text] [doi: 10.2196/16779] [Medline: 33090113]
14. Yanow D, Schwartz-Shea P, editors. Interpretation and Method. New York, NY: Routledge; 2014.
15. Weick KE, Sutcliffe KM, Obstfeld D. Organizing and the Process of Sensemaking. Organ Sci 2005 Aug;16(4):409-421. [doi: 10.1287/orsc.1050.0133]
16. BigMedilitycs. 2020. URL: https://www.bigmedilytics.eu/ [accessed 2020-06-26]
17. Bos A, Boselie P, Trappenburg M. Financial performance, employee well-being, and client well-being in for-profit and not-for-profit nursing homes: A systematic review. Health Care Manage Rev 2017;42(4):352-368. [doi: 10.1097/HMR.0000000000000121] [Medline: 28885990]
18. Van Elten HJ, Van der Kolk B, Sülz S. Do different uses of performance measurement systems in hospitals yield different outcomes? Health Care Manage Rev 2021 Sep 07;46(3):217-226. [doi: 10.1097/HMR.0000000000000261] [Medline: 31356357]
19. Berg M, Meijerink Y, Gras M, Goossensen A, Schellekens W, Haeck J, et al. Feasibility first: developing public performance indicators on patient safety and clinical effectiveness for Dutch hospitals. Health Policy 2005 Dec;75(1):59-73. [doi: 10.1016/j.healthpol.2005.02.007] [Medline: 16298229]
20. Wiener CL. Holding American hospitals accountable: rhetoric and reality. Nurs Inq 2004 Jun;11(2):82-90. [doi: 10.1111/j.1440-1800.2004.00212.x] [Medline: 15154887]
21. Freeman T. Using performance indicators to improve health care quality in the public sector: a review of the literature. Health Serv Manage Res 2002 May;15(2):126-137. [doi: 10.1258/0951484021912897] [Medline: 12028801]
22. Bal R. Playing the Indicator Game: Reflections on Strategies to Position an STS Group in a Multi-disciplinary Environment. Engaging STS 2017 Feb 17;3:41-52. [doi: 10.17351/ests2017.111]
23. Dahler-Larsen P. Constitutive Effects of Performance Indicators: Getting beyond unintended consequences. Public Manag Rev 2013 Apr 20;16(7):969-986. [doi: 10.1080/14719037.2013.770058]
24. Power M, Scheytt T, Soin K, Sahlin K. Reputational Risk as a Logic of Organizing in Late Modernity. Organ Stud 2009 Feb 01;30(2-3):301-324. [doi: 10.1177/0170840608101482]
25. Wallenburg I, Quartz J, Bal R. Making Hospitals Governable: Performativity and Institutional Work in Ranking Practices. Adm Soc 2016 Nov 30;51(4):637-663. [doi: 10.1177/0095399716680054]
26. Vosselman E. The 'performativity thesis' and its critics: Towards a relational ontology of management accounting. Account Bus Res 2013 Dec 16;44(2):181-203. [doi: 10.1080/00014788.2013.856748]
27. Objectives. BigMedylitics project website. 2020. URL: https://www.bigmedilytics.eu/big-data-project/#project-objectives [accessed 2020-01-09]
28. Defining and Measuring Productivity. OECD. URL: http://www.oecd.org/sdd/productivity-stats/40526851.pdf [accessed 2020-06-25]
29. Hollingsworth B. The measurement of efficiency and productivity of health care delivery. Health Econ 2008 Oct;17(10):1107-1128. [doi: 10.1002/hec.1391]
30. Djellal F, Gallouj F. The productivity challenge in services: measurement and strategic perspectives. Serv Ind J 2013 Feb;33(3-4):282-299. [doi: 10.1080/02642069.2013.747519]
31. Hamilton KES, Coates V, Kelly B, Boore JRP, Cundell JH, Gracey J, et al. Performance assessment in health care providers: a critical review of evidence and current practice. J Nurs Manag 2007 Nov;15(8):773-791. [doi: 10.1111/j.1365-2934.2007.00780.x] [Medline: 17944603]

32. Wiig S, Aase K, von Plessen C, Burnett S, Nunes F, Weggelaar AM, For QUASER-team. Talking about quality: exploring how 'quality' is conceptualized in European hospitals and healthcare systems. BMC Health Serv Res 2014 Oct 11;14:478 [FREE Full text] [doi: 10.1186/1472-6963-14-478] [Medline: 25303933]
33. Kaplan RS, Norton DP. The Balanced Scorecard - Measures that Drive Performance. Harv Bus Rev 1992(January-February):71-79.
34. Kaplan RS. Conceptual Foundations of the Balanced Scorecard. In: Chapman CS, Hopwood AG, Shields MD, editors. Handbooks of Management Accounting Research. Oxford, United Kingdom: Elsevier; 2009:1253-1269.
35. Chong CA, Nguyen GC, Wilcox ME. Trends in Canadian hospital standardised mortality ratios and palliative care coding 2004-2010: a retrospective database analysis. BMJ Open 2012;2(6):1-7 [FREE Full text] [doi: 10.1136/bmjopen-2012-001729] [Medline: 23131397]
36. Fernando SM, Neilipovitz D, Sarti AJ, Rosenberg E, Ishaq R, Thornton M, et al. Monitoring intensive care unit performance-impact of a novel individualised performance scorecard in critical care medicine: a mixed-methods study protocol. BMJ Open 2018 Jan 21;8(1):e019165 [FREE Full text] [doi: 10.1136/bmjopen-2017-019165] [Medline: 29358441]
37. Dixon-Woods M, Campbell A, Chang T, Martin G, Georgiadis A, Heney V, et al. A qualitative study of design stakeholders' views of developing and implementing a registry-based learning health system. Implement Sci 2020 Mar 06;15(1):16 [FREE Full text] [doi: 10.1186/s13012-020-0976-1] [Medline: 32143678]
38. Kleinsmann M. Understanding Collaborative Design. Delft, the Netherlands: Technical University Delft; 2006.
39. Clemensen J, Rothmann MJ, Smith AC, Caffery LJ, Danbjorg DB. Participatory design methods in telemedicine research. J Telemed Telecare 2017 Oct 01;23(9):780-785. [doi: 10.1177/1357633X16686747] [Medline: 28027678]
40. Fowler N, Lindahl M, Sköld D. The projectification of university research. Int J Manag Proj 2015 Jan 5;8(1):9-32. [doi: 10.1108/IJMPB-10-2013-0059]
41. Wehrens R, Oldenhof L, Bal R. On Staging Work: How Research Funding Bodies Create Adaptive Coherence in Times of Projectification. Science, Technology, & Human Values 2021 Apr 08:016224392110055. [doi: 10.1177/01622439211005557]
42. Felt U. Under the Shadow of Time: Where Indicators and Academic Values Meet. Engaging STS 2017 Feb 17;3:53-63. [doi: 10.17351/ests2017.109]
43. Torka M. Projectification of Doctoral Training? How Research Fields Respond to a New Funding Regime. Minerva 2018 Jan 17;56(1):59-83. [doi: 10.1007/s11024-018-9342-8]
44. Dixit RA, Hurst S, Adams KT, Boxley C, Lysen-Hendershot K, Bennett SS, et al. Rapid development of visualization dashboards to enhance situation awareness of COVID-19 telehealth initiatives at a multihospital healthcare system. J Am Med Inform Assoc 2020 Jul 01;27(9):1456-1461 [FREE Full text] [doi: 10.1093/jamia/ocaa161] [Medline: 32618999]
45. Lee M, Lin F, Chen S, Hsu W, Lin S, Chen T, et al. Web-Based Dashboard for the Interactive Visualization and Analysis of National Risk-Standardized Mortality Rates of Sepsis in the US. J Med Syst 2020 Jan 11;44(2):54. [doi: 10.1007/s10916-019-1509-9] [Medline: 31927706]
46. Weijers M, Bastiaenen C, Feron F, Schröder K. Designing a Personalized Health Dashboard: Interdisciplinary and Participatory Approach. JMIR Form Res 2021 Feb 09;5(2):e24061 [FREE Full text] [doi: 10.2196/24061] [Medline: 33560229]

Abbreviations

BSC: balanced scorecard
KPI: key performance indicator

Edited by A Mavragani; submitted 05.05.21; peer-reviewed by I de Loo, SD Boie, R Lee; comments to author 20.07.21; revised version received 10.09.21; accepted 15.12.21; published 22.02.22.

Please cite as:
van Elten HJ, Sülz S, van Raaij EM, Wehrens R
Big Data Health Care Innovations: Performance Dashboarding as a Process of Collective Sensemaking
J Med Internet Res 2022;24(2):e30201
URL: https://www.jmir.org/2022/2/e30201
doi:10.2196/30201
PMID:35191847

Viewpoint

Digital Health Paradox: International Policy Perspectives to Address Increased Health Inequalities for People Living With Disabilities

Robin van Kessel[1,2,3], MSc, PhD; Rok Hrzic[1*], MSc, MD; Ella O'Nuallain[4*], BCom; Elizabeth Weir[5*], MSc, PhD; Brian Li Han Wong[3,6,7], MSc; Michael Anderson[8], MSc, MD; Simon Baron-Cohen[5], PhD; Elias Mossialos[8,9], PhD

[1]Department of International Health, Care and Public Health Research Institute, Maastricht University, Maastricht, Netherlands

[2]Studio Europa, Maastricht University, Maastricht, Netherlands

[3]Global Health Workforce Network Youth Hub, World Health Organization, Geneva, Switzerland

[4]Public Sector Strategy Team, Deloitte Consulting Pty Ltd, Sydney, Australia

[5]Autism Research Center, Department of Psychiatry, University of Cambridge, Cambridge, United Kingdom

[6]The Lancet and Financial Times Commission on Governing Health Futures 2030: Growing up in a digital world, Global Health Centre, The Graduate Institute, Geneva, Switzerland

[7]Steering Committee, European Public Health Association Digital Health Section, Utrecht, Netherlands

[8]Department of Health Policy, London School of Economics and Political Science, London, United Kingdom

[9]Institute of Global Health Innovation, Imperial College London, London, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Robin van Kessel, MSc, PhD
Department of International Health
Care and Public Health Research Institute
Maastricht University
Minderbroedersberg 4-6
Maastricht, 6211LK
Netherlands
Phone: 31 43 388 2222
Email: r.vankessel@maastrichtuniversity.nl

Abstract

The COVID-19 pandemic accelerated the uptake of digital health worldwide and highlighted many benefits of these innovations. However, it also stressed the magnitude of inequalities regarding accessing digital health. Using a scoping review, this article explores the potential benefits of digital technologies for the global population, with particular reference to people living with disabilities, using the autism community as a case study. We ultimately explore policies in Sweden, Australia, Canada, Estonia, the United Kingdom, and the United States to learn how policies can lay an inclusive foundation for digital health systems. We conclude that digital health ecosystems should be designed with health equity at the forefront to avoid deepening existing health inequalities. We call for a more sophisticated understanding of digital health literacy to better assess the readiness to adopt digital health innovations. Finally, people living with disabilities should be positioned at the center of digital health policy and innovations to ensure they are not left behind.

(J Med Internet Res 2022;24(2):e33819) doi:10.2196/33819

KEYWORDS

digital health; eHealth; health policy; health systems; disability; inclusion; digital technologies; people living with disabilities

The Digital Paradox as a Modern Wicked Problem

As foundational parts of society continue to digitalize [1], the risk of existing inequalities worsening cannot be overstated. Digital divides constitute a wicked challenge in European countries [2], especially considering how digital literacy and access to digital infrastructure differ across age, sex, socioeconomic and educational strata, place of residence, and disability status [2,3]. The COVID-19 pandemic has highlighted the benefit of advances in digital health technologies [4], which could be deployed at a large scale to reach population groups that are otherwise difficult to reach. This reflects the paradox of digital health that we are currently facing: the potential that digital health innovations hold can be transformational for delivering care to underserved population groups, but these groups are most likely to be excluded from the digital world through their sociodemographic characteristics [5].

One example of such groups is the autism community. Recent research suggests that autistic individuals have comparatively shorter life spans [6,7] and are more likely to experience nearly any chronic physical and mental health condition [8-10]. Despite greater health care utilization and expenditure [10-12], autistic individuals report lower quality health care, worse health care access, and poorer patient-provider communication [11,13-17]. In qualitative interviews, autistic adults expressed difficulties evaluating or describing their health, sensory sensitivities, executive function, body awareness, slow processing speed, stigmatization, and systemic barriers [13-16]. As a result, the autism community can be considered highly vulnerable to digital exclusion, not only because of the aforementioned challenges but also due to the poorer educational and employment outcomes associated with autism [18].

The European Commission recognizes in the Union of Equality: Strategy for the Rights of Persons with Disabilities 2021-2030 that accelerated digital transformation offers opportunities to remotely deliver high-quality care to people living with disabilities that is tailored to their needs [19]. However, the precise actions needed to develop an inclusive digital health system that benefits people living with disabilities instead of excluding them, as well as best practices, are undefined.

It is important to note that this article takes a holistic approach to digital health. In other words, digital health as a concept can refer to a technology, user experience, service, product, process, infrastructure of its own, or part of the ecological system of health services [20-24]. It is important to be mindful of all these potential framings of digital health to encourage a systems thinking approach to how digital health may be incorporated in the status quo of health care. While there is great potential in the deployment of digital health services, the digital paradox makes it painfully clear that the utility of digital health is proportional to how well it meets the health care needs of the person interacting with it.

In this article, we explore the ways in which digitalization affects the health sector, specifically looking at the distribution of information, social determinants of health, and access to digital infrastructure. We then highlight how these digital developments may affect people living with disabilities, using the autism community as a case study. Afterward, we showcase existing policy perspectives from Sweden, Australia, Canada, Estonia, the United Kingdom, the United States, Singapore, Japan, and South Korea that either facilitate or obstruct the creation of digital health systems accessible for people living with disabilities. Finally, we identify principles of digital health that can underpin and facilitate the safe and inclusive development of a digital health ecosystem. We believe this article may be of particular interest to policy makers, health practitioners, patient groups, health funders, social workers, and other stakeholders in the field of health care. Details on the methodology and the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) flowchart can be found in Multimedia Appendix 1.

Impact of Digital Technologies on the Distribution of Information

While the relationship between digital technologies and health is nonlinear and multifaceted, digital technologies can facilitate the distribution and dissemination of public health information and action (eg, widespread broadcasts by the World Health Organization throughout the COVID-19 pandemic). Digital health can also enhance clinical and laboratory work by supporting administrative tasks, leaving more time for health care workers to spend with their patients [1,23]. This, however, requires a careful balance between increasing efficiencies through digital means and reducing the cognitive burden of frontline staff (explored later in this paper). Telemedicine can improve access to life-saving medical care. Soon, digital technology will augment clinical diagnostic processes, improve data collection and analysis, and provide targeted support in the form of precision medicine and precision public health [1,22,23,25].

On the other hand, health disinformation and misinformation can easily and rapidly be spread in a digital society through social media, contributing to the continued existence and strengthening of factually incorrect health beliefs [26]. A recent notable example of health misinformation has been the tremendous decline in vaccine confidence [27]. Machingaidze and Wiysonge [27] explain how digital technologies can be excellent tools for self-education, a key component of vaccination decision-making. However, they also present various challenges of using digital technologies at scale, many of which are present in high-income countries, including misinformation, incomplete information, and conflicting and complicated scientific information that may be difficult to understand [27].

Digital Technologies and Social Determinants of Health

Digital technology can also further entrench established social determinants of health [24]. For example, wearable technologies hold great potential to equip the general public with the ability to proactively monitor and manage their health care, fitness, and aging [28,29], although this comes at a financial cost.

Presently, this expense is borne by consumers with the ability to pay. This creates inequities in access to personalized health data, placing it at the fingertips of wealthier individuals and further widening health divides between socioeconomic groups. As such, when advocating for the widespread uptake of digital health among vulnerable groups, we must avoid constructing or fueling the paradigm of digital health as a pinnacle of health consumerism, as this would only expand the existing digital paradox [21]. A recent study in the United Kingdom highlights that access to free or affordable health care and individual behaviors and choices dominate the public perception of what impacts an individual's health and that 24% of the UK population believes health is entirely the responsibility of the individual [30]. As a result, the belief that healthier choices are available to more resourceful population groups is commonly propagated, disregarding the potential and complex effects of social determinants of health.

Furthermore, economies of scale achieved through increased use will drive down the price of wearable technology [31]. A case study in rural India demonstrated the benefits of wearables for chronically underserved rural and remote populations, showing improved health outcomes, including decreased readmission and mortality rates, by monitoring health data and improving preventive care [32]. Furthermore, wearables reduced the need for costly transport and frequent doctor visits as a patient's health could be observed at home and transmitted to their doctor in another locality [32]. This remote monitoring approach could prove hugely beneficial for people living with disabilities, as wearables are able to address various dimensions of disability, including physical, sensory, emotional, and intellectual conditions [33]. This approach could provide people living with disabilities with greater autonomy, augment their ability to live independently, and improve educational and employment outcomes [34].

Wearables have been shown to enable people to become informed about their health status and conditions, give patients a sense of agency over their health metrics, and improve health literacy [29,35]. Moreover, health care providers and governments can harness digital technology's aggregated information and insights to improve the quality of diagnosis and broader population health [29,36]. As such, digital divides threaten the equitable access of health care and public health services and fundamentally disadvantage the strata of the population that cannot pilot digital technologies [3,22,24]. Therefore, measuring digital health literacy is critical as it will help evaluate the readiness of a health system to implement digital health innovations, design targeted educational tools before implementation, and avoid costly attempts that were certain to fail from the start [37].

Differences in Access to Digital Technologies and Infrastructure

Considerable work has been done to measure the world population's ability to access digital services in terms of available infrastructure and digital literacy [3,22,38]. Nevertheless, these efforts are often limited to a binary measure of internet access and an overly simplified set of digital skills [3,38]. As such, they fall short by failing to acknowledge complex or domain-specific skills required to maximize the benefits of digital medical services. Simultaneously, while digital technologies are frequently designed to be used at the individual level, a system-level approach is required to fully understand the breadth and depth of how digital technologies can shift or augment a (public) health system [23].

A crucial part of this system-level approach is the ability of frontline health care workers to engage with and effectively use digital (health) technologies and platforms to streamline their workflows. With digital health technologies, the technology must be user-friendly to relieve the cognitive burden of health workers so they can focus on providing medical care [39]. Fractured communication and different systems across facilities and departments are often significant reasons for increased cognitive load [39]. As such, by optimally managing and communicating information through standardized systems, the potentials of digital health could be maximized to achieve the most significant benefit for health professionals [39].

Nurses are well placed to drive the digital health literacy agenda and upskill their colleagues. Moreover, as nurses often act as the primary care providers for rural or vulnerable populations, they would benefit from developing the knowledge and skills to implement and use telehealth technologies [40]. Strengthening their ability to use digital health technologies should therefore be an integral part of education for nurses and other health professions. A multimodal training approach (ie, didactics, patient simulations, practice immersions, and project work) has enabled students to develop the skills required to embrace telehealth and be comfortable using it [40]. Simultaneously, medical training continues to focus heavily on the quantitative side while underserving the arts and humanities of medicine [41], inherently limiting the extent to which medical professionals can advocate for the influence of various social determinants on health. However, it is the social determinants that often play a multifaceted role in whether a patient is able to access digital health services. In order for digital health care to become a mainstream mode of health care, medical education needs to be rebalanced with adequate coverage of arts and humanities of medicine and social medicine to train future health professionals when it is appropriate to deploy digital services, especially as long as the digital divide remains as extensive as it is currently. Additionally, social prescribing (ie, referring people to a range of local, nonclinical services) is increasingly becoming a core role of clinicians [42]. This could be further ameliorated with specialized training programs to ensure they have the right tools and referral pathways on hand to deploy it effectively.

Virtual medicine has significantly changed the way health care professionals deliver care. The COVID-19 pandemic has accelerated the development of telehealth services to the point where they have become a viable pathway for remotely delivering accessible, cost-effective, and quality care to patients [40,43]. Although videoconferencing is increasingly regarded as an adequate substitute for face-to-face consultations, network coverage shortages and lack of digital access or literacy among doctors and patients have meant that a proportion of remote primary care still takes place via telephone [44,45]. This change

in the modality of care provision is also expected to persist after the pandemic as users appreciate the convenience and accessibility afforded by telehealth services [40,46,47], especially in remote areas or for vulnerable populations who prefer or require access to health care closer to home. Digital health must, in turn, become more accessible, as must the incorporation of features that enable its acceptance and the reimbursement of this modality [46]. This is particularly important in resource-limited settings and low- and middle-income countries, where accessible pathways must be established. This will maximize the potential of telehealth to transform health care delivery for the wider global population [48].

Digital Technologies, Health, and Disability

While people living with disabilities cannot be taken as a monolith, they are on average older, poorer, and less likely to have a regular health care provider (and thus less likely to seek health care regularly) [49]. They are also more likely to experience multiple co-occurring conditions [49]. These factors pose additional challenges for developing and integrating the infrastructure of telemedicine into the lives of people living with disabilities [50]. However, digital health platforms offer an opportunity for improved engagement with populations with complex health needs who may receive poorer quality health care services, including those on the neurodiversity spectrum.

Telemedicine appointments allow patients to interact with health care providers at home [40,48,51], limiting exposure to overwhelming and unknown sensory environments while also accommodating alternative forms of communication (eg, audio-based or chat-based), which may be preferred by autistic adults [11]. Telemedicine and digital health may provide an online record of health care appointments (past and future), diagnoses, and prescriptions that patients can view at any time [51], which may afford greater autonomy for autistic patients or caregivers in managing their health care. Finally, online access to health care records for patients—particularly those with complex care needs—can provide reminders for essential tasks (eg, attending future follow-up appointments, collecting prescriptions) and offer opportunities for health care records to be corrected in the event of miscommunication between patients and providers.

Systematic reviews of pediatric and adult populations found that services delivered via telehealth were equally effective or superior means of providing health care to people living with disabilities, including those on the autism spectrum [52-54]. Furthermore, other forms of digital health like wearable technologies may provide novel opportunities for remote monitoring of people living with disabilities in distress situations or promote healthy lifestyle choices (eg, diet, exercise, and sleep) [29,51]. These tend to be significantly worse among autistic adults relative to neurotypical people and are associated with excess risk of chronic cardiovascular conditions [55]. Although we have explored the potential advantages of integrating telemedicine into health care systems for autistic adults, these benefits may also apply to those with other disabilities and the general population. Thus, telemedicine may be an efficient and cost-effective means of beginning to close the gaps in equitable access to high-quality health care for all.

Policy Perspectives and Practices

This section shows an overview of guidelines established in 3 countries to help facilitate the (digital) inclusion of people living with disabilities. Each set of principles shown in Textboxes 1, 2, and 3 [56-59] lays a foundation for a collaborative environment where all stakeholders are involved and empowered in the development and implementation of digital health, allowing for different vulnerable groups to be included during the development process to ensure their needs are met by the end product.

Sweden's digital health initiatives have been ongoing for 15 years at the time of writing with its current Vision for eHealth 2025 being adopted in 2016 (Textbox 1) [56]. The Vision acknowledges the opportunities that digital health brings for improving the welfare and well-being of vulnerable communities. It states that for these opportunities to be actualized digital health should be nondiscriminatory and respond to the needs of different groups. This is possible given the inherent characteristic of digital health to be deployed at an individual level.

Textbox 1. Strategy for implementing Vision for eHealth 2025 in Sweden [57].

Objectives:

- Involve the individual as cocreator
- Ensure that all relevant information is easily available when needed
- Guarantee that personal information is processed safely and securely
- Acknowledge that development and digital transformations go hand-in-hand

Fundamental principles:

- Regulations should (1) safeguard individual rights such as privacy, equality, patient safety, and accessibility; (2) enable the most optimal use of digital health innovations; and (3) stand the test of time and be specific enough to be applied
- Terminology and semantics should be used consistently across administrative systems to ensure and improve interoperability
- Acknowledging that a lot of standardization occurs at the European Union level, common and cross-sectoral standards (ie, procedures of application and development) are promoted to prevent national or sector-specific unique solutions

Textbox 2. National Disability Strategy–United Kingdom [58].

The governmental approach to disability is guided by 5 principles:

- Ensuring equity and fairness: people living with disabilities will be empowered by advocating for fairness and equality in opportunities, experiences, and outcomes
- Considering disability from the start: inclusive and accessible approaches will be embedded in government functioning to mitigate the creation of excluding experiences for people living with disabilities
- Supporting independent living: initiatives that support all people living with disabilities to have choice, control, and autonomy in their life will be actively encouraged
- Increasing participation: a diverse group of people living with disabilities will be included in the design, development, and delivery of products, services, and policies
- Delivering joined-up responses: work will occur across organizational boundaries to improve data and evidence gathering to better understand the complex issues that affect people living with disabilities

Textbox 3. Accessible Canada Act [59].

The purpose of this legislation is to benefit all persons, especially people living with disabilities, by removing existing barriers and preventing new barriers within Canada on or before January 1, 2040. It sets out 7 principles to guide the transformation process:

- Every person must be treated with dignity regardless of their disabilities
- Every person must have the same opportunity to make for themselves the lives that they are able and wish to have regardless of their disabilities
- Every person must have barrier-free access to full and equal participation in society, regardless of their disabilities
- Every person must have meaningful options and be free to make their own choices, with support if they desire, regardless of their disabilities
- Laws, policies, programs, services, and structures must take into account the disabilities of persons, the different ways that persons interact with their environments, and the multiple and intersecting forms of marginalization and discrimination faced by persons
- People living with disabilities must be involved in the development and design of laws, policies, programs, services, and structures
- The development and revision of accessibility standards and the making of regulations must be done with the objective of achieving the highest level of accessibility for people living with disabilities

This sentiment also arose from the United Kingdom Disability Survey 2021, which indicates that "44% of [people living with disabilities] who received [online] formal care said it made them feel more in control or much more in control of their lives" [58]. For this transformation to occur sustainably, strong principles are required to guide and facilitate the transformation process.

In terms of operationalizing the fundamental principles in Textboxes 1, 2, and 3, Australia and Canada showcase good policy practices. Australia adopted the Digital Transformation Strategy 2018-2025 in 2018, which is geared toward the digital transition of government services [60]. The strategy includes the Digital Service Standard criteria, which extensively outlines how developers and innovators can ensure that their services are usable by every person who needs them [61]. The guidelines distinguish 4 phases of progress (discovery, alpha, beta, and live), each containing between 4 and 15 requirements that need to be met before the next phase can begin. During the first 3 phases, innovators must prove that people from different backgrounds and people living with disabilities were involved in the development process and that the innovation is accessible for these groups. While the guidelines were designed for innovation in government services, the criteria are good practice guidelines for inclusive digital innovation in every sector, including health care. These guidelines can be further augmented by making funding contingent on adherence to such guidelines. The Accessible Technology Program in Canada is an example of this. Even though the program description does not offer extensive guidelines like the ones set out in Australia, its overarching aim is similar: "to develop innovative assistive and adaptive digital devices and technologies for persons with disabilities" [62].

Alternatively, the United Kingdom provides an extensive list of organizations that give support, advice, or information for each stage of the development life cycle for digital health and care products [63].

When considering active citizen engagement, we must consider what meaningful or equitable engagement looks like in health research. Literature on various frameworks for patient or citizen engagement is rich [64-67], although there is a common consensus among these studies that the frameworks themselves are lackluster and no specific research on the involvement of people living with disabilities or vulnerable groups has been carried out. As such, we are not in the position to establish concrete guidelines for meaningfully including people living with disabilities in the co-design process. That said, there is a consensus that meaningful engagement adheres to a number of principles (eg, the INVVOLVE principles: Invest in co-design; Needs assessment; Vision roles, responsibilities, and rewards; Validate participants; Organize interaction carefully; Lead the engagement; Value patient time and input; Evaluate and report [67]). When determining a method for meaningful engagement of people living with disabilities, it is good practice to consider the ability and preferences of the participants [68]. For instance,

workshops can be a meaningful way to invite and interact with a representative group of people living with disabilities throughout the design process [69] but may only attract certain clusters that can handle the sensory or cognitive input paired with attending workshops. Ratwani and colleagues [70] further elaborate on some good and bad practices regarding user-centered design processes in the development of electronic health records that can be particularly relevant for informing co-design methods for people living with disabilities.

Estonia is currently regarded as the frontier of how an integrated digital health system can function [71,72]. Estonia offers citizens the ability to access diagnostic services, consultations, prescription refills, and referrals online. The digital health system is built on the basis of an information society, which was initiated in 1992. Having regained independence from the Soviet Union in 1990, Estonia was able to rebuild and develop a society pointed toward the Digital Age without the remnants of past times. In contrast, most Western countries are run by long-established governments and sometimes even longer established ideologies, which may constrain the adoption of transformational innovations that significantly challenge the status quo [73].

For example, health care coverage and insurance, including telemedicine access, remains disjointed and variable for people living with disabilities in the United States. In 1990, the Americans With Disabilities Act (ADA) was passed to protect people living with disabilities against discrimination, specifically requiring that health care entities provide full and equal access for people living with disabilities. However, until the passage of the Affordable Care Act (colloquially known as Obamacare) in 2010, insurers could discriminate against people living with disabilities by charging higher premiums or denying health care coverage by capitalizing on an exemption in the ADA [74]. As the United States does not offer universal health care or health insurance schemes, service access for people living with disabilities presently depends on insurance type: Medicaid (37.7%), Medicare (27.1%), private insurance (36.1%), military benefits (6.0%), and uninsured (8.5%) [75], in turn providing patchwork telemedicine coverage. This is particularly true for Medicaid (the largest single provider of health insurance coverage for people living with disabilities), as each state runs their Medicaid system independently, individually interpreting federal mandates (like the ADA) and definitions of telemedicine. As of 2021, both Medicare and Medicaid universally offer some access to telemedicine services [76]; however, these programs are limited based on location, plan type, service type, provider type, and provider licensing state, and prescriptions may not be allowed via telemedicine. Further, only 10 state-run Medicaid programs offer all 3 major types of telemedicine to enrollees: live conferencing, remote patient monitoring, and store-and-forward electronic health records [76]. As it stands, this system (or lack thereof) leaves people living with disabilities either unable to access a wide range of telemedicine services or forced to reckon with significant system-level barriers, including deciphering individual insurer policies (and sometimes individual state insurance policies) to determine which types of telemedicine coverage can be reimbursed, much less whether their specific provider will be covered. Finally, it cannot be overstated how access issues compound when dealing with intersectional populations (including those experiencing poverty, homelessness, racism, and sexism along with disabilities), in turn providing greater disparities in health care coverage as well as greater opportunities for closing the gap via telemedicine.

Policy searches were also performed for South Korea, Japan, and Singapore. However, it appeared that digital transformation strategies were not yet developed in these countries, which is not only a contrast to the Western world but also to the fact that these countries are perceived as leaders in the field of digital transformation [77].

Limitations

Before accurate recommendations for policy and practice can be established, the circumstances under which this article was written need to be considered. The findings of this article should be interpreted as scoping and are not definitive due to the unsystematic nature of data collection. Our aim is to inspire more thorough research and initial/preliminary action in the discussed topics. Additionally, the quality of included sources has not been assessed, which reinforces the need to interpret the results carefully.

Moreover, the author team consists almost exclusively of young professionals who grew up during a period when digital technologies flourished and had access to digital technologies during their formative years. The author team also exclusively consists of people with a Western background (European/North American). Consequently, the values, interpretations, opinions, and recommendations in this piece—while rooted in scientific evidence—may be considered progressive and transformational when presented to population groups that have different relationships with digital technologies. That said, the demographic makeup of the author team can also be considered a strength of the article, as youth and young professionals have a strong stake in the future development of the health system.

Principles for Policy, Research, and Practice

The policy perspectives above point toward a strong need to work collaboratively across sectors. Technological, social, organizational, and political innovations that fundamentally reimagine health care delivery must be embraced with the aim to establish a new social contract that is fit for purpose for delivering high-quality digital health care to people living with disabilities [78,79]. At the same time, these policy perspectives highlight how digital health is currently framed in policy as a technology or a service more so than anything else. It is framed to be used by a consumer or a patient rather than be delivered by a health professional.

For effective implementation of digital (health) technologies, legal challenges, security breaches, regulatory concerns, and industry barriers must be addressed. Existing legislation, regulations, and policies have largely been written for face-to-face health care delivery, and security regulations vary widely across (and even within) territories [43,44]. When

considering the developmental trajectory of Estonia in the field of digital health, it becomes evident that a collaboration of public and private stakeholders, including national governments, patient groups, international governmental and nongovernmental organizations, citizen representatives, health care representatives, and social sectors, must convene to consider whether the values underpinning existing health care structure are still fit for purpose. Subsequently, this coalition should move to develop a high-level consensus on definitions, boundaries, protocols, monitoring, evaluation, and data privacy that is needed to meet the rapid advancement of digital health [47]. Specifically, they should focus on reimbursement of digital health care, privacy/cybersecurity, liability, licensure, and technology access, while ensuring that all countries participating in this coalition meet an internationally recognized minimum standard [80]. International and national policy makers need to amend the existing policy framework to support and enable new ways of working driven by digital health [57,60]. In the interim, given that many professional liability policies currently exclude digital health from coverage, health care practitioners are advised to acquire additional coverage to ensure protection from liability issues [43].

Designing an ecosystem of digital health technologies with health equity in mind is imperative to avoid deepening health inequalities [81]. In other words, the burden of implementing digital health innovations should not fall on citizens by focusing solely on upskilling after digital health tools are developed. This applies doubly when aiming to reach people living with disabilities and other groups that are frequently digitally excluded and may garner unique benefits from these services. Following the Swedish and Australian examples, policy makers can draft policies that include vulnerable groups as integral parts of the co-design process. Additionally, funding agencies can follow the Canadian example and require that people from different backgrounds, including people living with disabilities, are included at every stage to get funding for a project.

Future research should work toward adopting a broader and more sophisticated understanding of digital health literacy (for instance, the transactional model of eHealth literacy [82]). Subsequently, digital health research should meaningfully engage with vulnerable population groups to ensure that new digital tools meet their individual needs [57,60,62,83]. Specifically, a new relationship between qualitative and quantitative research fields needs to be established that considers qualitative and quantitative research complementary to each other [84]. Related to this is the need for software developers to partner with user experience and user interface experts in designing digital health technologies accessible to everyone [85] and reduce rather than add to the workload of health care professionals. The relationship between qualitative and quantitative research may also need to be reconsidered as new data streams become available through digital health innovations.

In summary, there is a role for national governments, actors from private sectors (eg, telecommunication, technology), international/global governing bodies (eg, World Health Organization, Organisation for Economic Co-operation and Development), and other stakeholders to collaboratively create a regulatory framework that ensures a high level of quality in digital health care through independent evaluation and certification [20]. The task of quality assurance should not fall to the individual. In other words, a systems thinking approach involving as many stakeholders as possible is paramount to safely and sustainably implementing digital transformations of this magnitude. This approach should be complemented by novel forms of collaborative governance and leadership that are goal-oriented, embrace risk-taking, and derive value from improved patient and societal outcomes rather than monetary gain [78]. The good practices and recommendations in this article have been condensed into domain-specific principles shown in Textbox 4. These principles are designed to facilitate the initial stages of development of a digital health ecosystem that is accessible to all.

Textbox 4. Principles focused on policy, practice, governance, and education for the sustainable integration of digital health innovations.

Health policy:

- Digital health is framed as a multidimensional concept that can refer to a technology, user experience, service, product, process, infrastructure of its own, or part of the ecological system of health services
- Digital health care services are a viable way to deliver health care to underserved populations that is comparable in quality to analog health care
- Health data, a major driver of digital health care, is stored safely and securely with the patient having autonomy over with whom the information is shared
- People living with disabilities and vulnerable groups are core subjects upon which digital health-related policy is built

Health care design and delivery:

- The design and development of digital health services are in close collaboration with people living with disabilities and other vulnerable population groups
- The involvement of people living with disabilities and other vulnerable groups in co-designing digital health services is systematic, purposeful, and equitable
- Clear agreements regarding expectations, input, and reward are made before the design process starts
- Digital health services are released and monitored collaboratively with people living with disabilities and other vulnerable groups

Governance:

- Governance approaches within health care embrace risk-taking, learn to function based on incomplete information, and assess value by outcome measures instead of monetary costs
- Collaborative governance is a central tenet in the governing process of health care that drives the other fields of policy making, health care design and delivery, and education

Education:

- A baseline course on the social determinants of health and their relevance to individual health is included in educational curricula of primary or secondary education
- A more in-depth course on the social determinants of health or social medicine is included in the educational curricula of all health professions
- Digital health literacy is a skill that needs to be maintained and updated intermittently to consistently engage with digital health services over time

Conclusions

Digital technology is fundamentally reshaping the way health care is delivered. The current scope of its use is not fully living up to its potential. The COVID-19 pandemic saw digital health catapulted into widespread use as virtual delivery became a priority for many health services worldwide. However, the pandemic also emphasized that digital health in its current form exacerbates health inequalities. At greatest risk of digital exclusion are older people, people in rural and remote areas, and people living with disabilities.

While the futuristic use of digital technologies is auspicious, we must allow both the policy landscapes and global digital health literacy levels to catch up with the rapid advances in technology [86]. We must ensure that digital technologies are used equitably and do not become an exclusive domain of high-income countries or populations. People living with disabilities and other vulnerable groups must be put at the center of digital health development on a global scale. Ultimately, achieving equitable access to digital health will significantly benefit the health and well-being of the wider population, especially vulnerable groups, and go a long way in reaching the Agenda for Sustainable Development 2030.

Acknowledgments

EO'N participated in this research independently of her role at Deloitte, where she is employed. Furthermore, Deloitte did not provide funding or other support for this research. EW received funding from the Corbin Charitable Trust, Autism Research Trust, Rosetrees Trust, and the Collaboration for Leadership in Applied Health Research and Care East of England at Cambridgeshire and Peterborough National Health Service (NHS) Foundation Trust. SBC received grant 214322\Z\18\Z from the Wellcome Trust. For the purpose of Open Access, the author has applied a CC-BY public copyright license to any author accepted manuscript version arising from this submission. In addition, SBC received grant 777394 from Innovative Medicines Initiative 2 Joint Undertaking (the funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results. Any views expressed are those of the authors and not necessarily those of the funders). The Innovative Medicines Initiative 2 Joint Undertaking receives support from the European Union's Horizon

2020 research and innovation program, the European Federation of Pharmaceutical Industries and Associations, AUTISM SPEAKS, Autistica, and the Simons Foundation Autism Research Initiative (SFARI). SBC also received funding from the Autism Research Trust, SFARI, the Templeton World Charitable Fund, the Medical Research Council, and the National Institute for Health Research (NIHR) Cambridge Biomedical Research Centre. The research was supported by the NIHR Collaboration for Leadership in Applied Health Research and Care East of England at Cambridgeshire and Peterborough National Health Service Foundation Trust. The views expressed are those of the authors and not necessarily those of the NHS, NIHR, or Department of Health and Social Care. The funding agencies did not play a role at any stage in the development of this article.

Authors' Contributions

RVK, EM, and RH were in charge of conceptualization. RVK, RH, EO'N, EW, and BLHW were responsible for data collection, writing, reviewing, and editing the manuscript. MA, SBC, and EM were involved in the reviewing and editing of the manuscript. RVK supervised the project. RH, EO'N, and EW contributed equally to this manuscript and retain the right to list themselves as the second author for the purpose of their CV and grant applications.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary materials.
[DOCX File , 220 KB - jmir_v24i2e33819_app1.docx]

References

1. Policy@Manchester. Analysis and ideas on addressing digital inequalities. Manchester: University of Manchester URL: https://documents.manchester.ac.uk/display.aspx?DocID=55160&utm [accessed 2022-02-01]
2. Scheerder A, van Deursen A, van Dijk J. Determinants of Internet skills, uses and outcomes. A systematic review of the second- and third-level digital divide. Telematics Informatics 2017 Dec;34(8):1607-1624. [doi: 10.1016/j.tele.2017.07.007]
3. van Kessel R, Wong BLH, Rubinić I, Czabanowska K. Is Europe prepared to go digital? Making the case for developing digital capacity: an exploratory analysis of Eurostat survey data. PLOS Digital Health (in press) 2022:1. [doi: 10.1371/journal.pdig.0000013]
4. Carter P, Anderson M, Mossialos E. Health system, public health, and economic implications of managing COVID-19 from a cardiovascular perspective. Eur Heart J 2020 Jul 14;41(27):2516-2518 [FREE Full text] [doi: 10.1093/eurheartj/ehaa342] [Medline: 32320040]
5. Sanders R. Digital inclusion, exclusion and participation. Glasgow: Iriss; 2020 Apr 09. URL: https://www.iriss.org.uk/resources/esss-outlines/digital-inclusion-exclusion-and-participation [accessed 2022-02-01]
6. Hirvikoski T, Mittendorfer-Rutz E, Boman M, Larsson H, Lichtenstein P, Bölte S. Premature mortality in autism spectrum disorder. Br J Psychiatry 2016 Mar;208(3):232-238. [doi: 10.1192/bjp.bp.114.160192] [Medline: 26541693]
7. Hwang YI, Srasuebkul P, Foley K, Arnold S, Trollor JN. Mortality and cause of death of Australians on the autism spectrum. Autism Res 2019 May 25;12(5):806-815. [doi: 10.1002/aur.2086] [Medline: 30802364]
8. Croen LA, Zerbo O, Qian Y, Massolo ML, Rich S, Sidney S, et al. The health status of adults on the autism spectrum. Autism 2015 Oct 24;19(7):814-823. [doi: 10.1177/1362361315577517] [Medline: 25911091]
9. Weir E, Allison C, Warrier V, Baron-Cohen S. Increased prevalence of non-communicable physical health conditions among autistic adults. Autism 2021 Apr;25(3):681-694 [FREE Full text] [doi: 10.1177/1362361320953652] [Medline: 32907337]
10. Weiss JA, Isaacs B, Diepstra H, Wilton AS, Brown HK, McGarry C, et al. Health concerns and health service utilization in a population cohort of young adults with autism spectrum disorder. J Autism Dev Disord 2018 Jan;48(1):36-44 [FREE Full text] [doi: 10.1007/s10803-017-3292-0] [Medline: 28900771]
11. Nicolaidis C, Raymaker D, McDonald K, Dern S, Boisclair WC, Ashkenazy E, et al. Comparison of healthcare experiences in autistic and non-autistic adults: a cross-sectional online survey facilitated by an academic-community partnership. J Gen Intern Med 2013 Jun;28(6):761-769 [FREE Full text] [doi: 10.1007/s11606-012-2262-7] [Medline: 23179969]
12. Zerbo O, Qian Y, Ray T, Sidney S, Rich S, Massolo M, et al. Health care service utilization and cost among adults with autism spectrum disorders in a U.S. integrated health care system. Autism Adulthood 2019 Mar;1(1):27-36. [doi: 10.1089/aut.2018.0004]
13. Camm-Crosbie L, Bradley L, Shaw R, Baron-Cohen S, Cassidy S. 'People like me don't get support': autistic adults' experiences of support and treatment for mental health difficulties, self-injury and suicidality. Autism 2019 Aug 29;23(6):1431-1441 [FREE Full text] [doi: 10.1177/1362361318816053] [Medline: 30497279]
14. Crane L, Adams F, Harper G, Welch J, Pellicano E. "Something needs to change": mental health experiences of young autistic adults in England. Autism 2019;23(2):477-493. [doi: 10.1177/1362361318757048]

15. Mason D, Ingham B, Urbanowicz A, Michael C, Birtles H, Woodbury-Smith M, et al. A systematic review of what barriers and facilitators prevent and enable physical healthcare services access for autistic adults. J Autism Dev Disord 2019 Aug;49(8):3387-3400 [FREE Full text] [doi: 10.1007/s10803-019-04049-2] [Medline: 31124030]
16. Nicolaidis C, Raymaker DM, Ashkenazy E, McDonald KE, Dern S, Baggs AE, et al. "Respect the way I need to communicate with you": healthcare experiences of adults on the autism spectrum. Autism 2015 Oct;19(7):824-831 [FREE Full text] [doi: 10.1177/1362361315576221] [Medline: 25882392]
17. Raymaker D, McDonald K, Ashkenazy E, Gerrity M, Baggs A, Kripke C, et al. Barriers to healthcare: instrument development and comparison between autistic adults and adults with and without other disabilities. Autism 2017 Nov;21(8):972-984 [FREE Full text] [doi: 10.1177/1362361316661261] [Medline: 27663266]
18. Lord C, Elsabbagh M, Baird G, Veenstra-Vanderweele J. Autism spectrum disorder. Lancet 2018 Aug;392(10146):508-520. [doi: 10.1016/s0140-6736(18)31129-2]
19. European C. Union of equality: strategy for the rights of persons with disabilities 2021-2030. Luxembourg: Directorate General for Employment, Social Affairs, and Inclusion URL: https://data.europa.eu/doi/10.2767/31633 [accessed 2021-09-07]
20. Mathews SC, McShea MJ, Hanley CL, Ravitz A, Labrique AB, Cohen AB. Reply: some considerations on digital health validation. NPJ Digit Med 2019 Oct 17;2(1):103 [FREE Full text] [doi: 10.1038/s41746-019-0176-z] [Medline: 31637310]
21. Cohen AB, Mathews SC, Dorsey ER, Bates DW, Safavi K. Direct-to-consumer digital health. Lancet Digit Health 2020 Apr;2(4):e163-e165 [FREE Full text] [doi: 10.1016/S2589-7500(20)30057-1] [Medline: 33328077]
22. Honeyman M, Maguire D, Evans H, Davies A. Digital technology and health inequalities: a scoping review. Cardiff: Public Health Wales NHS Trust; 2020. URL: https://phw.nhs.wales/publications/publications1/digital-technology-and-health-inequalities-a-scoping-review/ [accessed 2022-02-01]
23. Budd J, Miller BS, Manning EM, Lampos V, Zhuang M, Edelstein M, et al. Digital technologies in the public-health response to COVID-19. Nat Med 2020 Aug 07;26(8):1183-1192. [doi: 10.1038/s41591-020-1011-4] [Medline: 32770165]
24. Sieck CJ, Sheon A, Ancker JS, Castek J, Callahan B, Siefer A. Digital inclusion as a social determinant of health. NPJ Digit Med 2021 Mar 17;4(1):52 [FREE Full text] [doi: 10.1038/s41746-021-00413-8] [Medline: 33731887]
25. Murray CJL, Alamro NMS, Hwang H, Lee U. Digital public health and COVID-19. Lancet Public Health 2020 Sep;5(9):e469-e470 [FREE Full text] [doi: 10.1016/S2468-2667(20)30187-0] [Medline: 32791051]
26. Chan C, Sounderajah V, Daniels E, Acharya A, Clarke J, Yalamanchili S, et al. The reliability and quality of youtube videos as a source of public health information regarding COVID-19 vaccination: cross-sectional study. JMIR Public Health Surveill 2021 Jul 08;7(7):e29942 [FREE Full text] [doi: 10.2196/29942] [Medline: 34081599]
27. Machingaidze S, Wiysonge CS. Understanding COVID-19 vaccine hesitancy. Nat Med 2021 Aug 16;27(8):1338-1339. [doi: 10.1038/s41591-021-01459-7] [Medline: 34272500]
28. Sultan N. Reflective thoughts on the potential and challenges of wearable technology for healthcare provision and medical education. Int J Inf Manag 2015 Oct;35(5):521-526. [doi: 10.1016/j.ijinfomgt.2015.04.010]
29. 5 Health Benefits of Wearable Tech. Welbi. 2017 Oct 25. URL: https://www.welbi.co/blog/5-health-benefits-of-wearable-tech [accessed 2021-08-16]
30. Elwell-Sutton T, Marshall L, Bibby J, Volmert A. Reframing the conversation on social determinants. London: Health Foundation/Framework Institute; 2019. URL: https://www.health.org.uk/sites/default/files/upload/publications/2019/Reframing%20the%20conversation%20on%20social%20determinants.pdf [accessed 2022-02-01]
31. Godfrey A, Hetherington V, Shum H, Bonato P, Lovell N, Stuart S. From A to Z: wearable technology explained. Maturitas 2018 Jul;113:40-47 [FREE Full text] [doi: 10.1016/j.maturitas.2018.04.012] [Medline: 29903647]
32. Jhajharia S, Pal D, Verma D. Wearable computing and its application. Int J Comput Sci Inf Technol 2014;5(4):5700-5704 [FREE Full text]
33. Anderson C, Anderson K. Wearable technology: meeting the needs of individuals with disabilities and its applications to education. In: Buchem I, Klamma R, Wild F, editors. Perspectives on Wearable Enhanced Learning (WELL). Cham: Springer International Publishing; 2019:59-77.
34. Moon N, Baker P, Goughnour K. Designing wearable technologies for users with disabilities: accessibility, usability, and connectivity factors. J Rehab Assistive Technol Eng 2019 Aug 13;6:205566831986213. [doi: 10.1177/2055668319862137]
35. Wu M, Luo J. Wearable technology applications in healthcare: a literature review. Online J Nurs Inform 2019 Nov 25;23(3):1 [FREE Full text]
36. Chau KY, Lam MHS, Cheung ML, Tso EKH, Flint SW, Broom DR, et al. Smart technology for healthcare: exploring the antecedents of adoption intention of healthcare wearable technology. Health Psychol Res 2019 Mar 11;7(1):8099. [doi: 10.4081/hpr.2019.8099] [Medline: 31583292]
37. Pfob A, Sidey-Gibbons C, Schuessler M, Lu S, Xu C, Dubsky P, et al. Contrast of digital and health literacy between IT and health care specialists highlights the importance of multidisciplinary teams for digital health—a pilot study. JCO Clinical Cancer Informatics 2021 Jun(5):734-745. [doi: 10.1200/cci.21.00032]
38. Measuring Digital Development: Facts and Figures. Geneva: International Telecommunications Union; 2020. URL: https://www.itu.int/en/ITU-D/Statistics/Pages/ff2020interactive.aspx [accessed 2022-02-01]
39. Collins R. Cognitive overload on the front line of healthcare. 2021. URL: https://www.talkinghealthtech.com/podcast/140-cognitive-overload-on-the-front-line-of-healthcare-dr-rhonda-collins-vocera [accessed 2021-08-16]

40. Rutledge CM, Kott K, Schweickert PA, Poston R, Fowler C, Haney TS. Telehealth and eHealth in nurse practitioner training: current perspectives. Adv Med Educ Pract 2017;8:399-409 [FREE Full text] [doi: 10.2147/AMEP.S116071] [Medline: 28721113]
41. Moniz T, Golafshani M, Gaspar CM, Adams NE, Haidet P, Sukhera J, et al. The prism model: advancing a theory of practice for arts and humanities in medical education. Perspect Med Educ 2021 Aug 29;10(4):207-214 [FREE Full text] [doi: 10.1007/s40037-021-00661-0] [Medline: 33914287]
42. Royal Australian College of General Practitioners, Consumers Health Forum of Australia. Social Prescribing Roundtable. 2019 Nov. URL: https://chf.org.au/sites/default/files/social_prescribing_roundable_report_chf_racgp_v11.pdf [accessed 2022-02-01]
43. Balestra M. Telehealth and legal implications for nurse practitioners. J Nurse Pract 2018 Jan;14(1):33-39. [doi: 10.1016/j.nurpra.2017.10.003]
44. Kichloo A, Albosta M, Dettloff K, Wani F, El-Amir Z, Singh J, et al. Telemedicine, the current COVID-19 pandemic and the future: a narrative review and perspectives moving forward in the USA. Fam Med Community Health 2020 Aug;8(3):e000530 [FREE Full text] [doi: 10.1136/fmch-2020-000530] [Medline: 32816942]
45. Shaw SE, Seuren LM, Wherton J, Cameron D, A'Court C, Vijayaraghavan S, et al. Video consultations between patients and clinicians in diabetes, cancer, and heart failure services: linguistic ethnographic study of video-mediated interaction. J Med Internet Res 2020 May 11;22(5):e18378 [FREE Full text] [doi: 10.2196/18378] [Medline: 32391799]
46. Kruse CS, Krowski N, Rodriguez B, Tran L, Vela J, Brooks M. Telehealth and patient satisfaction: a systematic review and narrative analysis. BMJ Open 2017 Aug 03;7(8):e016242 [FREE Full text] [doi: 10.1136/bmjopen-2017-016242] [Medline: 28775188]
47. Fisk M, Livingstone A, Pit SW. Telehealth in the context of COVID-19: changing perspectives in Australia, the United Kingdom, and the United States. J Med Internet Res 2020 Jun 09;22(6):e19264 [FREE Full text] [doi: 10.2196/19264] [Medline: 32463377]
48. Doraiswamy S, Abraham A, Mamtani R, Cheema S. Use of telehealth during the COVID-19 pandemic: scoping review. J Med Internet Res 2020 Dec 01;22(12):e24087. [doi: 10.2196/24087] [Medline: 33147166]
49. Disability Impacts All of Us Infographic. Atlanta: Centers for Disease Control and Prevention; 2019. URL: https://www.cdc.gov/ncbddd/disabilityandhealth/infographic-disability-impacts-all.html [accessed 2021-08-17]
50. Annaswamy TM, Verduzco-Gutierrez M, Frieden L. Telemedicine barriers and challenges for persons with disabilities: COVID-19 and beyond. Disabil Health J 2020 Oct;13(4):100973 [FREE Full text] [doi: 10.1016/j.dhjo.2020.100973] [Medline: 32703737]
51. Silva BMC, Rodrigues JJPC, de la Torre Díez I, López-Coronado M, Saleem K. Mobile-health: a review of current state in 2015. J Biomed Inform 2015 Aug;56:265-272 [FREE Full text] [doi: 10.1016/j.jbi.2015.06.003] [Medline: 26071682]
52. Camden C, Pratte G, Fallon F, Couture M, Berbari J, Tousignant M. Diversity of practices in telerehabilitation for children with disabilities and effective intervention characteristics: results from a systematic review. Disabil Rehabil 2019 Apr 12:1-13. [doi: 10.1080/09638288.2019.1595750] [Medline: 30978110]
53. Ellison KS, Guidry J, Picou P, Adenuga P, Davis TE. Telehealth and autism prior to and in the age of COVID-19: a systematic and critical review of the last decade. Clin Child Fam Psychol Rev 2021 Sep;24(3):599-630 [FREE Full text] [doi: 10.1007/s10567-021-00358-0] [Medline: 34114135]
54. Sutherland R, Trembath D, Roberts J. Telehealth and autism: a systematic search and review of the literature. Int J Speech Lang Pathol 2018 Jun;20(3):324-336. [doi: 10.1080/17549507.2018.1465123] [Medline: 29709201]
55. Weir E, Allison C, Ong KK, Baron-Cohen S. An investigation of the diet, exercise, sleep, BMI, and health outcomes of autistic adults. Mol Autism 2021 May 08;12(1):31 [FREE Full text] [doi: 10.1186/s13229-021-00441-x] [Medline: 33964967]
56. Swedish Ministry of Health and Social Affairs, Swedish Association of Local Authorities and Regions. Vision for eHealth 2025. 2016 Aug 19. URL: https://www.government.se/4a3e02/contentassets/b0fd09051c6c4af59c8e33a3e71fff24/vision-for-ehealth-2025.pdf [accessed 2022-02-01]
57. Swedish Ministry of Health and Social Affairs, Swedish Association of Local Authorities and Regions. A strategy for implementing Vision for eHealth 2025: the next steps. 2020 Feb. URL: https://ehalsa2025.se/wp-content/uploads/2021/02/Strategy-2020-2022_eng.pdf [accessed 2022-02-01]
58. National Disability Strategy. Government of the United Kingdom. 2021 Jul. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1006098/National-Disability-Strategy_web-accesible-pdf [accessed 2022-02-01]
59. Accessible Canada Act. Canadian Ministry of Justice. 2019. URL: https://laws-lois.justice.gc.ca/eng/acts/A-0.6/ [accessed 2021-09-07]
60. Digital Transformation Strategy 2018-2025. Australian Government Digital Transformation Agency. URL: https://dta-www-drupal-20180130215411153400000001.s3.ap-southeast-2.amazonaws.com/s3fs-public/files/digital-transformation-strategy/digital-transformation-strategy.pdf [accessed 2022-02-01]
61. Digital Service Standard criteria. Australian Government Digital Transformation Agency. 2020. URL: https://www.dta.gov.au/help-and-advice/digital-service-standard/digital-service-standard-criteria [accessed 2021-08-30]

62. Accessible Technology Program. Government of Canada. 2017. URL: http://www.ic.gc.ca/eic/site/118.nsf/eng/home [accessed 2021-08-30]
63. Digital health and care. Government of the United Kingdom. 2018. URL: https://www.gov.uk/government/publications/uk-life-sciences-support/digital-health-and-care [accessed 2021-09-07]
64. Jull JE, Davidson L, Dungan R, Nguyen T, Woodward KP, Graham ID. A review and synthesis of frameworks for engagement in health research to identify concepts of knowledge user engagement. BMC Med Res Methodol 2019 Nov 21;19(1):211 [FREE Full text] [doi: 10.1186/s12874-019-0838-1] [Medline: 31752691]
65. Banner D, Bains M, Carroll S, Kandola DK, Rolfe DE, Wong C, et al. Patient and public engagement in integrated knowledge translation research: are we there yet? Res Involv Engagem 2019;5(1):1 [FREE Full text] [doi: 10.1186/s40900-019-0139-1] [Medline: 30805202]
66. Phoenix M, Nguyen T, Gentles SJ, VanderKaay S, Cross A, Nguyen L. Using qualitative research perspectives to inform patient engagement in research. Res Involv Engagem 2018 Jul 6;4(1):20 [FREE Full text] [doi: 10.1186/s40900-018-0107-1] [Medline: 30002874]
67. Slattery P, Saeri AK, Bragge P. Research co-design in health: a rapid overview of reviews. Health Res Policy Syst 2020 Feb 11;18(1):17 [FREE Full text] [doi: 10.1186/s12961-020-0528-9] [Medline: 32046728]
68. Maun R, Fabri M, Trevorrow P. Adapting participatory design activities for autistic adults: a review. Des User Exp Usability Des Divers Well Soc Dev 2021:300-314. [doi: 10.1007/978-3-030-78224-5_21]
69. Fabri M, Andrews P. Human-centered design with autistic university students: interface, interaction and information preferences. In: Marcus A, editor. Design, User Experience, and Usability: Novel User Experiences. Cham: Springer International Publishing; 2016:157-166.
70. Ratwani RM, Fairbanks RJ, Hettinger AZ, Benda NC. Electronic health record usability: analysis of the user-centered design processes of eleven electronic health record vendors. J Am Med Inform Assoc 2015 Nov;22(6):1179-1182. [doi: 10.1093/jamia/ocv050] [Medline: 26049532]
71. Mill R. Developing an integrated e-health system in Estonia. Tallinn: Estonian e-Health Foundation; 2016 Jun. URL: https://www.integratedcare4people.org/media/files/CaseProfileEstonia.pdf [accessed 2022-02-01]
72. Estonia and WHO to work together on digital health and innovation. Geneva: World Health Organization; 2020. URL: https://www.euro.who.int/en/countries/estonia/news/news/2020/10/estonia-and-who-to-work-together-on-digital-health-and-innovation [accessed 2021-08-30]
73. Sachs J. The Ages of Globalization: Geography, Technology, and Institutions. New York: Columbia University Press; 2020.
74. Blake V. Rethinking the Americans with Disabilities Act's Insurance Safe Harbor. Laws 2017 Nov 07;6(4):25. [doi: 10.3390/laws6040025]
75. Kennedy J, Wood EG, Frieden L. Disparities in insurance coverage, health services use, and access following implementation of the Affordable Care Act: a comparison of disabled and nondisabled working-age adults. Inq J Health Care Organ Provis Financ 2017 Jan 01;54:46958017734031 [FREE Full text] [doi: 10.1177/0046958017734031] [Medline: 29166812]
76. Feke T. Medicare and Medicaid increase your telehealth options. New York: Verywell Health; 2021 Sep 02. URL: https://www.verywellhealth.com/medicare-and-medicaid-coverage-for-telehealth-4682549 [accessed 2021-10-17]
77. E-Government Development Index Country Data 2020. United Nations Division for Public Institutions and Digital Government. 2021. URL: https://publicadministration.un.org/egovkb/en-us/Data-Center [accessed 2021-09-08]
78. Mazzucato M. Mission Economy: A Moonshot Guide to Changing Capitalism. New York: Harper Business; 2021.
79. Shafik M. What We Owe Each Other: A New Social Contract for a Better Society. Princeton: Princeton University Press; 2021.
80. Hoffman D. Increasing access to care: telehealth during COVID-19. J Law Biosci 2020;7(1):1 [FREE Full text] [doi: 10.1093/jlb/lsaa043] [Medline: 32843985]
81. Wong WF, LaVeist TA, Sharfstein JM. Achieving health equity by design. JAMA 2015 Apr 14;313(14):1417-1418. [doi: 10.1001/jama.2015.2434] [Medline: 25751310]
82. Paige SR, Stellefson M, Krieger JL, Anderson-Lewis C, Cheong J, Stopka C. Proposing a transactional model of ehealth literacy: concept analysis. J Med Internet Res 2018 Oct 02;20(10):e10175 [FREE Full text] [doi: 10.2196/10175] [Medline: 30279155]
83. Morton E, Barnes SJ, Michalak EE. Participatory digital health research: a new paradigm for mHealth tool development. Gen Hosp Psychiatry 2020 Sep;66:67-69 [FREE Full text] [doi: 10.1016/j.genhosppsych.2020.07.005] [Medline: 32702489]
84. Tanweer A, Gade E, Krafft P, Dreier S. Why the data revolution needs qualitative thinking. Harv Data Sci Rev 2021 Jul:1. [doi: 10.1162/99608f92.eee0b0da]
85. Mota L, David R, Botelho D, Oliveira V. UX Research applied to digital transformation of primary care in Brazil. Eur J Public Health 2020 Sep 01;30(Supplement_5):ckaa166.070. [doi: 10.1093/eurpub/ckaa166.070]
86. Wong B, Maaß L, Vodden A, van Kessel R, Sorbello S, Buttigieg S, European Public Health Association (EUPHA) Digital Health Section. The dawn of digital public health in Europe: Implications for public health policy and practice. Lancet Reg Health Eur 2022 Mar;14:100316 [FREE Full text] [doi: 10.1016/j.lanepe.2022.100316] [Medline: 35132399]

Abbreviations

ADA: Americans With Disabilities Act
INVVOLVE principles: Invest in co-design; Needs assessment; Vision roles, responsibilities, and rewards; Validate participants; Organize interaction carefully; Lead the engagement; Value patient time and input; Evaluate and report
NIHR: National Institute for Health Research
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-analyses
SFARI: Simons Foundation Autism Research Initiative

Edited by R Kukafka; submitted 24.09.21; peer-reviewed by A Garcia, M Gotsis; comments to author 16.10.21; revised version received 26.10.21; accepted 15.11.21; published 22.02.22.

Please cite as:
van Kessel R, Hrzic R, O'Nuallain E, Weir E, Wong BLH, Anderson M, Baron-Cohen S, Mossialos E
Digital Health Paradox: International Policy Perspectives to Address Increased Health Inequalities for People Living With Disabilities
J Med Internet Res 2022;24(2):e33819
URL: https://www.jmir.org/2022/2/e33819
doi:10.2196/33819
PMID:35191848

Original Paper

The Effect of Adjunct Telephone Support on Adherence and Outcomes of the Reboot Online Pain Management Program: Randomized Controlled Trial

Tania Gardner[1,2], PhD; Regina Schultz[3], PhD; Hila Haskelberg[1], PhD; Jill M Newby[4], PhD; Jane Wheatley[2], DPsych; Michael Millard[1], MBBS; Steven G Faux[2,4], MBBS; Christine T Shiner[1,2,4], PhD

[1]Clinical Research Unit for Anxiety and Depression, St Vincent's Hospital, Sydney, Australia

[2]Department of Pain Medicine, St Vincent's Hospital, Sydney, Australia

[3]Royal Ryde Rehabilitation Hospital, Sydney, Australia

[4]University of New South Wales, Sydney, NSW, Australia

Corresponding Author:
Tania Gardner, PhD
Clinical Research Unit for Anxiety and Depression, St Vincent's Hospital
406 Victoria St
Sydney, 2010
Australia
Phone: 61 0410449766
Email: taniagardner@optusnet.com.au

Abstract

Background: Internet-based treatment programs present a solution for providing access to pain management for those unable to access clinic-based multidisciplinary pain programs. Attrition from internet interventions is a common issue. Clinician-supported guidance can be an important feature in web-based interventions; however, the optimal level of therapist guidance and expertise required to improve adherence remains unclear.

Objective: The aim of this study is to evaluate whether augmenting the existing Reboot Online program with telephone support by a clinician improves program adherence and effectiveness compared with the web-based program alone.

Methods: A 2-armed, CONSORT (Consolidated Standards of Reporting Trials)–compliant, registered randomized controlled trial with one-to-one group allocation was conducted. It compared a web-based multidisciplinary pain management program, Reboot Online, combined with telephone support (n=44) with Reboot Online alone (n=45) as the control group. Participants were recruited through web-based social media and the This Way Up service provider network. The primary outcome for this study was adherence to the Reboot Online program. Adherence was quantified through three metrics: completion of the program, the number of participants who enrolled into the program, and the number of participants who commenced the program. Data on adherence were collected automatically through the This Way Up platform. Secondary measures of clinical effectiveness were also collected.

Results: Reboot Online combined with telephone support had a positive effect on enrollment and commencement of the program compared with Reboot Online without telephone support. Significantly more participants from the Reboot Online plus telephone support group enrolled (41/44, 93%) into the course than those from the control group (35/45, 78%; χ^2_1=4.2; *P*=.04). Furthermore, more participants from the intervention group commenced the course than those from the control group (40/44, 91% vs 27/45, 60%, respectively; χ^2_1=11.4; *P*=.001). Of the participants enrolled in the intervention group, 43% (19/44) completed the course, and of those in the control group, 31% (14/45) completed the course. When considering the subgroup of those who commenced the program, there was no significant difference between the proportions of people who completed all 8 lessons in the intervention (19/40, 48%) and control groups (14/27, 52%; χ^2_1=1.3; *P*=.24). The treatment efficacy on clinical outcome measures did not differ between the intervention and control groups.

Conclusions: Telephone support improves participants' registration, program commencement, and engagement in the early phase of the internet intervention; however, it did not seem to have an impact on overall course completion or efficacy.

Trial Registration: Australian New Zealand Clinical Trials Registry ACTRN12619001076167; https://anzctr.org.au/Trial/Registration/TrialReview.aspx?ACTRN=12619001076167

(J Med Internet Res 2022;24(2):e30880) doi:10.2196/30880

KEYWORDS

chronic pain; online pain management; telephone support; clinician guidance; adherence; digital health; eHealth; internet interventions; multidisciplinary

Introduction

Background

The global prevalence of chronic pain is estimated to be 10% to 33% [1-3], and pain conditions continue to be a leading cause of worldwide disability [4]. The accepted best practice for managing chronic pain is a multidisciplinary pain program that involves attending a group program—typically convened face to face at a hospital or clinical site—facilitated by members of a multidisciplinary team [5].

A main challenge to the management of chronic pain is the lack of services available within existing health care systems. Access to multidisciplinary pain management groups can be compromised for people living in rural areas, those with family or work commitments, and those with mental health issues that limit their ability to engage and be involved in a social group environment [6]. Furthermore, because of the COVID-19 global pandemic, increased social isolation and reduced accessibility of face-to-face clinical interactions have further compromised access to required pain management services for many people [7]. Web-based treatment programs present a solution for providing access to pain management for those unable to access clinic-based multidisciplinary pain programs [6,7] and for delivering care at reduced personnel costs [7-9].

The evidence to date suggests that internet-based health interventions that provide a combination of information and some form of user support—such as decision management support, behavior change support, or social support—lead to improvements in knowledge as well as behavioral and clinical outcomes [8,10-13]. Generally, internet-based pain programs focus on unimodal interventions that offer a single discipline of therapy, most commonly psychological therapies such as internet-delivered cognitive behavioral therapy (CBT) [14]. These programs typically show small to moderate effects on pain-related disability and interference as well as pain catastrophizing and are a cost-effective alternative to face-to-face therapy [13,14].

Reboot Online was one of the first internet-based pain management programs that offered a multidisciplinary model. The program has been shown to significantly improve pain self-efficacy and reduce pain-related disability, kinesiophobia, and psychological distress compared with usual care [15,16]. These studies were carried out within a strict clinical trial paradigm. Although highly effective for those participants who completed the program, adherence rates were moderate, with only 61% of the trial participants completing all the lessons. Adherence (completion of all 8 lessons of the program) and effectiveness (significant changes in clinical outcome measures) of the program in real-world conditions were subsequently evaluated [17]. Without the regulated framework and rigor embedded within a clinical trial design, the adherence to the Reboot Online program reduced to 41% when used under routine care conditions.

Attrition, or dropout of study participants, is a common problem for internet-based treatments [18,19]. In a systematic review of internet interventions for chronic pain [19], the range in attrition levels across published trials was considerable, from 4% to 54%. Different methods have been used to prevent dropout, such as telephone support, personalized reminders and feedback, and financial incentives [19]. However, it remains unclear how helpful these methods are [19].

Clinician-supported guidance can be an important feature in web-based interventions for mental health disorders; however, the optimal level and pattern of therapist guidance required to improve adherence remains unclear [20]. Having human social support, whether it be peer-led or specialized clinician–led, has been suggested to help support healthy behavior change [21,22] and has been shown to have small effects on improving adherence compared with unguided interventions for individuals with chronic pain [22-25]. The addition of clinician-supported guidance has not yet been tested in routine care using the Reboot Online program.

Objectives

We aim to evaluate whether augmenting the existing Reboot Online program with telephone support by a clinician improves program adherence and effectiveness compared with the web-based program alone. We hypothesized that the Reboot Online plus telephone support group would have better adherence (defined as higher number of enrollments, higher number of commencements, and higher completion rates) to the program than the control group who received the Reboot Online program only.

Methods

Study Design

The study was a 2-armed, CONSORT (Consolidated Standards of Reporting Trials)–compliant, registered randomized controlled trial with one-to-one group allocation. It compared a web-based multidisciplinary pain management program, *Reboot Online*, combined with telephone support with *Reboot Online* alone as the control group.

The study was approved by the human research ethics committee of St Vincent's Hospital, Sydney, Australia (2019/ETH08682). The trial was prospectively registered on the Australian New

Zealand Clinical Trials Registry (ACTRN12619001076167), and the protocol was followed as per the registry.

Participants

Participants were recruited between September 2019 and April 2020 through social media advertisements (eg, Facebook and Twitter) and through the This Way Up service provider network. To replicate real-world conditions, the trial was run through the established web-based platform through which the Reboot Online program is available: *This Way Up* [26]. Applicants completed a 2-step screening process. All applicants first completed web-based screening questionnaires about their chronic pain symptoms and demographic details, followed by a telephone interview to confirm their chronic pain diagnosis and study eligibility. The Mini International Neuropsychiatric Interview (version 5.0.0) [27] for major depressive disorder (MDD) and risk assessment modules were also administered during the telephone interview as a screening and diagnostic tool for depression. MDD helps us to compare the sample characteristics with our previous trials and establish the number of individuals who had comorbid depression, which also assists in clinical support. MDD plus risk assessment modules help establish how the clinician will prioritize and triage calls as well as monitor and ensure that participants who are at risk are appropriately supported throughout the program.

Applicants were eligible for inclusion if they (1) had experienced pain for >3 months, (2) were aged ≥18 years, (3) were a resident of Australia, (4) had their pain assessed by a physician in the past 3 months, (5) were prepared to provide contact details of their general practitioner, (6) had access to the internet and computer or tablet and deemed themselves to have basic computer literacy, and (7) were fluent in written and spoken English. Applicants were excluded if they had psychosis, bipolar disorder, were actively suicidal (eg, demonstrated intent or a plan or had a recent suicide attempt). These exclusion criteria were included to be consistent with other trial exclusion criteria and to be able to adequately evaluate participants' needs and assist their access to more intensive clinical support if required. Participants were also excluded if they had participated in a group-based pain management program within the last 6 months so that their prior treatment or learning would not confound the results. Participants were also ineligible for the study if they had surgical procedures scheduled in the 6 months after their application; had commenced or made significant changes to their psychotropic medication in the previous 2 months before intake assessment; or had commenced face-to-face CBT within 4 weeks of the intake assessment. These exclusion criteria were included because of their potential confounds. We wanted to explore the effects of clinician support on program adherence and outcomes and minimize the confounding effects of these features.

Participants provided electronic informed consent before being enrolled in the study; they also provided permission for their pain physician (or general practitioner) to be notified by the study team detailing their involvement in the trial. Measures of psychological distress (Kessler-10 Psychological Distress Scale [28,29]) were monitored throughout the program for any risk of harm.

Randomization

The eligible participants were randomly assigned to either the intervention or control group in a 1:1 allocation based on a random number sequence generated [30]. An individual web-based link for baseline data questionnaires through REDCap (Research Electronic Data Capture [31]) was sent to the eligible participants to complete. Once the initial baseline questionnaires were completed, an individual registration link to Reboot Online through the usual *This Way Up* process was sent to the participants, which gave them access to register for the course, complete pretreatment questionnaires, and commence the first lesson. Telephone screening interviewers at baseline remained blinded to the participants' group allocation. Study staff members conducting the telephone support intervention could not be blinded to group allocation. Group allocation was indicated to the participant by whether they received telephone support after registration into the program.

Intervention

Reboot Online Program (Control)

Participants received free access to the standard Reboot Online program [32]. Details of the content and design of the program have been previously outlined [16]. In brief, Reboot Online consists of 8 lessons delivered over 16 weeks. Participants follow an illustrated story of a fictional character who learns to self-manage her chronic pain using a multidisciplinary approach. The content delivers psychoeducation on the socio-psycho-biomedical nature of chronic pain within a multidisciplinary framework. The core lessons are combined with a *movement station* section: a graded, patient-led exercise program focusing on activity and exercise reactivation. This is combined with pacing: an activity-management technique aimed to maximize a person's activity by developing a structured, slow, and gradual planned increase in activity levels over a period of time. The level of activity is dependent on the planned quantity rather than pain intensity experienced by the patient. The program also incorporates SMART (specific, measurable, achievable, relevant, and timely) goal-setting principles, whereby the patient leads the process of identifying issues or problems that matter most to them, sets the goal that they want to accomplish, and develops strategies and goals that are specific, measurable, achievable, relevant, and within a time frame. This is coupled with evidence-based CBT skills activities, including thought challenging, activity planning, problem solving, effective communication, and flare-up management. Participants also have access to expert educational videos as well as tai chi and relaxation audio guides. See Table 1 for details of the lesson content of the program. All participants received automatic email communication to notify them when a lesson was available and to encourage lesson completion as well as engaging in activity.

Table 1. Reboot Online lesson content.

Lesson title	Lesson content	Homework activities	Lesson resources (PDF or video)
Lesson 1: What Is Chronic Pain and What Is the Best Way to Manage It?	• Chronic pain explained • The Brick Wall of Chronic Pain • Medication overview • The movement station	• Review of acute vs chronic pain • Complete own Brick Wall of Chronic Pain • Instructions for the movement station	• Welcome (video) • Medication management (video)
Lesson 2: Goal-Setting and Moving Toward Acceptance	• Scans, test results, and pain • The cycle of chronic pain • Moving toward acceptance • SMART[a] goals	• Reviewing the cycle of pain • Identifying acceptance, change, and goals • Setting short- and long-term SMART goals	• Medical imaging (video) • Making life changes (PDF)
Lesson 3: Movement, Pacing, and Daily Activity Scheduling	• Exploring the relationship between pain and activity • Learning about the "Boom Bust" pattern • Fear avoidance beliefs • Importance of pacing • Introduction to daily activity scheduling	• Identifying healthy vs unhealthy coping skills • Pacing activity • Keeping a movement diary	• Daily activity scheduling (PDF)
Lesson 4: Monitoring Thoughts and Recognizing Unhelpful Thinking Patterns	• Revision of chronic pain cycle and the link among thoughts, feelings, and behaviors • Recognizing unhelpful thinking patterns • Monitoring thoughts and learning to think more helpful thoughts	• Ways to recognize own unhelpful thinking patterns • Complete thought record • Continue to monitor and track thoughts	—[b]
Lesson 5: Mood and Pain, Thought Challenging, and Managing Arousal	• Challenging unhelpful thinking • Activity planning in practice • Emotions and pain • Strategies to manage anxiety • Mood	• Thought-challenging situations • Activity planning and monitoring • Managing anger: looking out for triggers and learning strategies • Practicing relaxation	• My Thought Challenging Worksheet (PDF)
Lesson 6: Stress Management and Getting Better Sleep	• What is stress? • How to manage stress better • Using problem solving • Barriers to good sleep with chronic pain and ways to get better sleep	• Recognizing sources of stress and own signs and symptoms • Structured problem-solving task • Sleep diary to improve habits	• Good sleep guide (PDF) • Better sleep for chronic pain (video)
Lesson 7: Communication and Relationships	• What is good communication? • Communication styles • Relationships with others • Ways to improve those relationships	• Communication skills practice • Task: select a relationship and explore how you would like that relationship to be different • Revisiting thought-challenging exercise	• Conversation skills tips (PDF) • Chronic pain and information for family and friends (PDF)
Lesson 8: Managing Flare-ups and Continuing Management of Chronic Pain	• Exploring flare-ups, relapse, and the need for a flare-up plan • Continuing chronic pain management • How to get further help	• Ten things to help during a relapse • Flare-up prevention plan • Summary	• What is a pain clinic? (video) • Congratulations (video)

[a]SMART: specific, measurable, achievable, relevant, and timely.

[b]Not available.

Reboot Online Program With Telephone Support (Intervention)

Participants received free access to the standard Reboot Online program, as outlined in the control intervention, in conjunction with telephone support. They received a phone call every fortnight for the duration of the program (maximum of 8 phone calls). There was a set maximum of 3 attempts if contact by phone was unsuccessful. The phone call was conducted by a single clinician (senior allied health clinician) experienced in the management of chronic pain. To replicate what may occur

in routine clinical care, we sought a flexible approach whereby the duration of the phone call was led by the needs of the patient. During the call, the participants were advised that the call was to check in and to see how their program was going. They were asked to report on their progress and encouraged to continue engaging with the program. The participants were also given the opportunity to discuss any challenges or hurdles they were experiencing, to problem-solve possible strategies to overcome them, and to receive feedback on their progress.

For both treatment groups, a senior clinician (senior pain physiotherapist TG and pain psychologist JW) experienced in chronic pain management monitored questionnaire responses, and if any responses indicated deterioration in well-being, phone contact was made with the participant and further clinical intervention was advised if indicated.

Outcome Measures

Participants were assessed using a suite of outcome measures, collected at three time points: baseline (immediately before commencing treatment), after treatment (weeks 16-17) and follow-up (3 months after completing posttreatment questionnaires).

The primary outcome for this study was adherence to the Reboot Online program. Adherence was quantified through three metrics: (1) completion of the program assessed through the number of participants who completed all 8 lessons of the intervention and who were classified as *completers*, whereas those who did not complete all 8 lessons were classified as *noncompleters*; (2) the number of participants who enrolled to undertake the program; and (3) the number of participants who commenced at least one lesson of the program. The rate of overall program completion was the primary outcome measure for adherence. These adherence measures were chosen to capture three key potential points where patients may drop out: (1) enrolling after being prescribed the course, (2) after completing baseline questionnaires and commencing the course, and (3) once engaged in the program after each lesson. Data on adherence were collected automatically through the This Way Up platform.

The secondary outcomes were as follows:

1. Pain Self-Efficacy Questionnaire [33] to assess participants' confidence to perform activities while in pain, with higher scores indicating greater confidence in functional capacity. A minimally clinically important difference (MCID) was considered to be 5.5 [34].
2. Tampa Scale for Kinesiophobia [35] to measure fear and avoidance of movement. An MCID was considered to be 6 [36].
3. Brief Pain Inventory [37] to assess pain severity and its impact on function through its 2 subscales for severity and pain interference. An MCID was considered to be 2 [38].
4. Pain Disability Index [39] to assess the degree to which chronic pain interferes with participants' daily activities and essential life activities. An MCID was considered to be between 8.5 and 9.5 [28].
5. International Physical Activity Questionnaire [29,40] to measure self-reported physical activity in the previous 7 days. Metabolic equivalents and level of physical activity were calculated using a preformed Microsoft Excel program [41]. This program calculated metabolic equivalents and levels of physical activity as low, moderate, and high according to International Physical Activity Questionnaire short form categorical scoring.
6. Kessler-10 Psychological Distress Scale [42,43] to measure psychological distress.

Self-reported data on clinical outcome measures were collected through the This Way Up platform as well as the REDCap tool hosted at St Vincent's Hospital, Sydney [44,45]. REDCap is a secure, web-based software platform designed to support data capture for research studies, providing (1) an intuitive interface for validated data capture, (2) audit trails for tracking data manipulation and export procedures, (3) automated export procedures for seamless data downloads to common statistical packages, and (4) procedures for data integration and interoperability with external sources.

Telephone Support Time and Satisfaction

Descriptive data pertaining to the amount of clinician contact and time spent conducting the telephone support calls were collected for all participants in the intervention group. Descriptive data on treatment satisfaction and perceived usefulness of the intervention were collected from all participants (both groups) during the posttreatment assessments. Additional user feedback was collected from the intervention group regarding how helpful and motivating they found the telephone support calls received as part of the intervention. Data on subjective participant feedback were quantitative in nature (collected on a rating scale from 1 to 10) and were analyzed descriptively through median values and IQRs (because of nonparametric data distribution).

Statistical Methods

Before commencement of the study, a power calculation was conducted to determine the minimum sample size required to detect a difference in the proportion of participants adhering to the course between the intervention and control groups. This was based on previous investigations [16], where 60% of the participants were observed to have adhered to the complete program under *control* conditions and an increase of 25% in adherence was predicted for the telephone support intervention group. Assuming a power of 80% and a Cronbach α set at .05, a minimum sample size of 39 participants in each group was needed to detect a 25% difference in the proportion of completers. To account for an anticipated approximate 10% dropout rate, a total of 44 participants were required in each study arm.

Descriptive statistics were used to characterize the demographic and clinical features of the study groups at baseline. Data on the time taken to deliver the telephone support and subjective participant feedback on the telephone support and web-based aspects of the Reboot Online program were also analyzed descriptively.

Cross-tabulations were used to examine between-group differences in the primary outcome of adherence. Chi-square analyses were performed with binary factors of group

(intervention and control) and adherence (yes or no) for each of the identified adherence metrics (enrollment, course commencement, and course completion). A Mann–Whitney *U* test was used to compare the median number of lessons completed by the intervention and control groups.

To examine treatment efficacy, intention-to-treat linear mixed models were implemented separately for each of the secondary outcome measures. This form of analysis can robustly account for the unbalanced nature of repeated measures data, including missing data because of participant dropout. Each model included group, time, and a group-by-time interaction as fixed factors. A random effect of participant was also included in each model, with an identity covariance structure used to model the covariance structure of the random intercept. Model fit indices supported the selection of an unstructured covariance matrix for each of the outcome measures. Significant treatment effects were followed with pair-wise comparisons of their estimated marginal means. For each outcome, the estimated marginal means derived from the model were used to calculate within-group effect sizes (Hedges g, adjusted for the correlation among time points). Pre- to posttreatment effect sizes (Hedges g, adjusted for the correlation among time points) were calculated, corresponding to the changes between pre- and posttreatment values and between posttreatment and follow-up values. Between-group effect sizes (Hedges g, adjusted for the correlation among time points) for each outcome after treatment and at follow-up were also calculated. Effect sizes of <0.49, 0.50-0.79, and >0.80 were considered to be small, moderate, and large, respectively [46].

Statistical analyses were conducted using SPSS software (version 25; IBM Corp), and results were considered significant where $P<.05$.

Results

Baseline Sample Characteristics

Baseline and sample characteristics are outlined in Table 2. Demographic data were available for all participants (N=89) who met the inclusion criteria. The participants had a mean age of 49.3 years (SD 16.1; range 21-86 years), and most of them were women (59/89, 66%). At baseline, 42% (37/89) of the participants reported taking simple analgesia for pain (nonopioid analgesics such as paracetamol and nonsteroidal anti-inflammatory drugs), 27% (24/89) took gabanoids, 43% (38/89) took antidepressant medication, and 44% (39/89) took opioid analgesia. Approximately 1 in 3 participants (28/89, 32%) met the criteria for MDD at the time of study enrollment.

Table 2. Baseline demographics (N=89).

Characteristics	Total sample (N=89)	Intervention (n=44)	Control (n=45)	Between-group comparison	
				Values	*P* value
Age (years), mean (SD)	49.3 (16.1)	48.0 (15.8)	50.6 (16.4)	t_{84}=–0.74	.46
Gender, n (%)				χ^2_2=2.2	.34
Male	28 (32)	13 (30)	15 (33)		
Female	59 (66)	29 (66)	30 (67)		
Nonbinary	2 (2)	2 (5)	0 (0)		
Rural status, n (%)				χ^2_1=0.9	.17
Major city	52 (58)	29 (66)	23 (51)		
Regional or remote	37 (42)	15 (34)	22 (49)		
Major depressive disorder, n (%)	28 (32)	16 (36)	12 (27)	χ^2_1=1.0	.33
Opioid analgesia, n (%)	39 (44)	14 (33)	25 (56)	χ^2_1=4.7	*.03* [a]
Antidepressants, n (%)	38 (43)	20 (47)	18 (40)	χ^2_1=0.4	.54
Anticonvulsants, n (%)	24 (27)	13 (30)	11 (24)	χ^2_1=0.4	.54
Benzodiazepines, n (%)	16 (18)	8 (19)	8 (18)	χ^2_1=0.0	.92
Simple analgesia, n (%)	44 (49)	18(41)	26 (58)	χ^2_1=2.5	.11

[a]Italicized results are significant at $P<.05$.

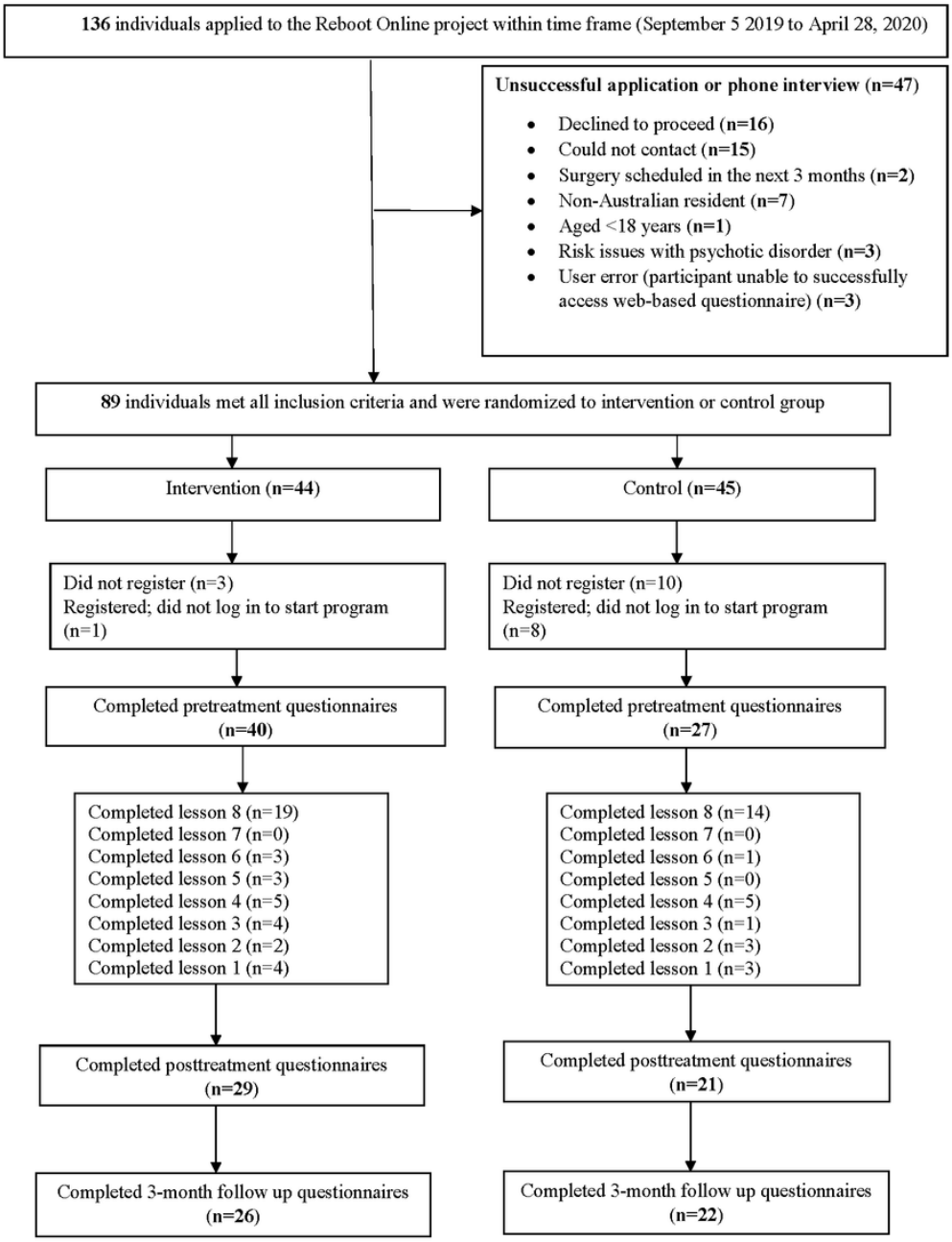
Adherence

Participant flow through the trial is illustrated in Figure 1. Reboot Online combined with telephone support had a positive effect on enrollment and commencement of the program compared with Reboot Online without telephone support. Significantly more participants from the Reboot Online plus telephone support group enrolled (41/44, 93%) into the course than those from the control group (35/45, 78%; χ^2_1=4.2; P=.04). Furthermore, more participants from the intervention group

commenced the course than those from the control group (40/44, 91% vs 27/45, 60%, respectively; χ^2_1=11.4; *P*=.001).

The median number of lessons completed by those in the intervention versus control groups was significantly different; the intervention group completed a median of 6 (IQR 3-8) lessons compared with a median of 2 (IQR 0-8) lessons in the control group (Mann–Whitney *U*=682; *P*=.009). More participants from the intervention group completed at least half the course (≥4 lessons; χ^2_1=5.0; *P*=.03) than those from the control group. The overall completion rate for both groups was 37% (33/89). Of the participants enrolled in the intervention group, 43% (19/44) completed the course, and of those in the control group, 31% (14/45) completed the course. When considering the subgroup of those who commenced the program, the overall course completion was 49% (33/67) from both groups; however, there was no significant difference between the proportions of people who completed all 8 lessons in the intervention (19/40, 48%) versus control groups (14/27, 52%; χ^2_1=1.4; *P*=.24).

Figure 1. Participant flowchart.



Effectiveness in Clinical Outcomes

Estimated marginal means at pretreatment, posttreatment, and 3-month follow-up as well as within- and between-group effect sizes for all outcome measures, are shown in Table 3. Significant main effects of time showing improvements from before to after treatment were observed for measures of pain self-efficacy ($F_{2,48}$=16.4; *P*<.001), kinesiophobia ($F_{2,52}$=10.3; *P*<.001), pain

interference ($F_{2,53}$=12.4; P<.001), pain disability ($F_{2,50}$=10.8; P<.001), and psychological distress ($F_{2,46}$=9.6; P<.001). There was no significant main effect of time with measures of pain severity ($F_{2,53}$=2.0; P=.15) or physical activity ($F_{2,55}$=1.8; P=.19). Within each group, posttreatment improvements were maintained at 3-month follow-up assessment, with no significant changes in outcomes between posttreatment and follow-up values.

There was no significant group-by-time interaction observed for any outcome measure, indicating that treatment efficacy on these outcome measures did not differ between the intervention and control groups.

Table 3. Estimated marginal means (EMMs) before and after treatment and at follow-up and within- and between-group effect sizes.

Outcome	Before treatment, EMM (SD)	After treatment, EMM (SD)	Follow-up, EMM (SD)	Pre- to posttreatment within-group comparison		Posttreatment to follow-up within-group comparison		Posttreatment between-group comparison, between-group effect size, Hedges g (95% CI)	Follow-up, between-group comparison, between-group effect size, Hedges g (95% CI)
				r	Within-group effect size, Hedges g (95% CI)	*r*	Within-group effect size, Hedges g (95% CI)		
PSEQ[a]									
Intervention	26.9 (10.9)	33.9 (12.0)	33.8 (12.3)	0.73	0.61 (0.08 to 1.14)[b]	0.92	0.01 (–0.57 to 0.59)	0.43 (–0.12 to –0.99)	0.32 (–0.26 to –0.89)
Control	28.7 (10.0)	38.9 (10.5)	37.7 (11.9)	0.37	0.98 (0.28 to 1.67)[b]	0.62	0.10 (–0.55 to 0.76)	N/A[c]	N/A
TSK[d]									
Intervention	37.4 (6.2)	33.1 (8.0)	33.2 (7.1)	0.60	0.60 (0.07 to 1.12)[b]	0.74	0.01 (–0.57 to 0.58)	0.50 (–0.05 to –1.04)	0.23 (–0.33 to –0.80)
Control	40.6 (5.9)	36.9 (7.2)	34.8 (7.1)	0.38	0.54 (–0.11 to –1.19)[e]	0.72	0.29 (–0.35 to 0.93)	N/A	N/A
BPI[f] severity									
Intervention	5.4 (1.4)	5.1 (1.5)	5.3 (1.4)	0.63	0.25 (–0.27 to –0.76)	0.81	0.19 (–0.39 to 0.77)	0.04 (–0.50 to –0.57)	0.36 (–0.21 to –0.94)
Control	5.4 (1.3)	5.0 (1.4)	4.8 (1.4)	0.66	0.28 (–0.36 to 0.91)	0.80	0.13 (–0.50 to 0.77)	N/A	N/A
BPI interference									
Intervention	6.4 (1.8)	5.3 (2.1)	4.9 (2.2)	0.75	0.59 (0.06 to 1.11)[b]	0.78	0.15 (–0.43 to 0.73)	0.01 (–0.5 to –0.55)	0.21 (–0.36 to 0.78)
Control	6.2 (1.7)	5.2 (1.9)	4.5 (2.2)	0.22	0.53 (–0.12 to –1.18)	0.75	0.36 (–0.28 to 1.00)	N/A	N/A
PDI[g]									
Intervention	40.1 (12.2)	30.9 (15.6)	31.0 (16.2)	0.69	0.65 (0.12 to 1.18)[b]	0.63	0.01 (–0.57 to 0.59)	0.07 (–0.47 to –0.60)	0.23 (–0.34 to –0.80)
Control	38.2 (11.5)	31.9 (13.9)	27.3 (15.8)	0.52	0.49 (–0.16 to –1.13)[e]	0.82	0.31 (–0.33 to 0.94)	N/A	N/A
IPAQ[h]									
Intervention	1571.9 (1709.8)	2783.4 (2790.3)	2353.7 (2391.5)	0.68	0.53 (0.01 to 1.05)[e]	0.56	0.17 (–0.41 to 0.75)	0.18 (–0.37 to –0.74)	0.24 (–0.33 to –0.81)
Control	2176.6 (1472.6)	2278.5 (2698.4)	1766.6 (2525.9)	0.49	0.05 (–0.54 to –0.64)	0.70	0.19 (–0.40 to 0.78)	N/A	N/A
K10[i]									
Intervention	27.3 (6.8)	23.2 (7.5)	23.6 (8.7)	0.76	0.57 (0.04 to 1.09)[b]	0.71	0.05 (–0.53 to 0.63)	0.07 (–0.49 to –0.63)	0.02 (–0.54 to –0.59)
Control	26.4 (5.9)	22.7 (6.3)	23.4 (8.3)	0.40	0.59 (–0.12 to –1.30)[e]	0.71	0.09 (–0.58 to 0.77)	N/A	N/A

[a]PSEQ: Pain Self-Efficacy Questionnaire.

[b]Significance at $P<.001$.

[c]N/A: not applicable.

[d]TSK: Tampa Scale for Kinesiophobia.

[e]Significance at $P<.05$.

[f]BPI: Brief Pain Inventory.

[g]PDI: Pain Disability Index.

[h]IPAQ: International Physical Activity Questionnaire.

[i]K10: Kessler-10 Psychological Distress Scale.

Telephone Support Sessions

Over the course of the study, all participants in the telephone support group (n=44) had at least one telephone support session and 30% (13/44) had 8 support sessions, with a mean number of 5.7 (SD 2.4) sessions for each participant. The total cumulative telephone time for the intervention group was 3240 minutes, which comprised multiple attempts to make telephone contact as well as successful telephone support sessions, and the overall active telephone time in which the clinician was engaged with the participant was 1780 (54.94%) minutes. This equated to a mean total of 73.6 (SD 31.8) minutes of clinician time per participant over the entire program, of which 40.5 (SD 26.8) minutes was active time spent engaging with the participant. The mean clinician time at each active session was 7.9 (SD 5.5) minutes. No additional time was spent with any participants in either group for follow-up because of acute deterioration in well-being.

Treatment Satisfaction

When asked to rate their satisfaction with the treatment received in this study on a scale from 0 to 10, the participants provided a median rating of 7 (IQR 5-8). Furthermore, when asked to rate how helpful they found the study intervention for managing their chronic pain, the participants responded with a median rating of 8 (IQR 5-9) out of 10. There was no significant difference between the groups for scores of either satisfaction or how helpful they found the Reboot Online program. Of the 44 participants in the intervention group, 35 (79%) reported that the telephone support made it easier for them to participate in or complete the Reboot Online program. When asked how motivating or encouraging they found the telephone support aspect of their treatment specifically (rated out of 10), the median score of the participants was 7 (IQR 5-9).

Discussion

Principal Findings

This study aimed to investigate whether additional clinician guidance, in the form of telephone support, would significantly enhance adherence to the Reboot Online program for chronic pain and improve clinical outcomes. We found that Reboot Online combined with telephone coaching resulted in improved rates of enrollment and commencement (onboarding) of the program compared with the usual Reboot Online program. The overall completion rate from both groups was 37% (33/89; intervention 19/44, 43%, and control 14/45, 31%). The overall completion rate for those who commenced the program from both groups was 49% (33/67); however, there was no significant difference in overall completion rate between the groups (intervention 19/40, 48%; control 14/27, 52%) once the program was commenced. These completion rates are consistent with our real-world adherence outcome of 41% [17] and reflect published attrition rates: 4% to 54% [20]. Furthermore, there was no significant group difference in treatment efficacy on a variety of outcome measures.

The most effective way to provide clinician guidance adjunct to other interventions remains unclear. Little is known as to whether providing guidance evenly timed over the entire course of an intervention is most effective or whether initial support gradually tapered to self-management would be more effective [18]. Our results showed a significant difference between the groups in the rates of enrollment and commencement of the program—the *onboarding phase*. Once participants commenced the program, there was no significant difference in completion rates between the groups, regardless of phone support. This would suggest that telephone support is most effective in improving the onboarding of persons with chronic pain undertaking web-based treatments. The average amount of time involved for each telephone session (including time spent calling to contact participants together with time of consult) was 9.2 (SD 6.0) minutes, with a mean active consult time of 7.9 (SD 5.5) minutes. Thus, the telephone support intervention did not seem onerous and therefore could be easily incorporated into a clinical caseload. When considering where to best use clinical resources, clinician time may be best used in the early phase of onboarding to maximize adherence.

The telephone support provided in this study was delivered by experienced allied health clinicians with specialized training in pain management. The content of the telephone support calls was basic and consisted mainly of reaffirmation of the course material as well as feedback on progress and suggestions to address any barriers that the participants were experiencing. There was no specialized psychological treatment or intervention to address mood or cognition delivered through the calls themselves; nor was there detailed clinical education or advice on movement and physical activity. As the support did not provide specialized intervention, in future it could be provided by a trained peer or junior clinician to streamline the provision of clinical resources, which may be adequate for patients who are not clinically complex. Other studies have shown little difference between clinician and peer-led coaching.

It has been suggested that telephone support provides the possibility to make a diagnosis, tailor the intervention, and actively assist patients to access other needed services [47]. A triage system could also be used to identify persons with more complex issues who need to be referred to specialist clinician intervention. This may be relevant for those with chronic pain who are particularly afraid of movement because of fear of pain because it has been shown that high kinesiophobia negatively

effects adherence rates [17]. Research investigating the level of telephone support stratified according to kinesiophobia scores may be worthwhile, with this study being underpowered to investigate this aspect.

The act of individually tailoring a web-based intervention has been suggested to help with adherence and outcomes [47]. Standardized and nontailored internet treatments may leave little room for patient and clinician preferences [48] regarding course content, format, and learning styles. Chien et al [49] identified 5 engagement subtypes in patients enrolled in an internet-delivered CBT program. Their study suggests that by identifying subtypes of patient engagement, programs can be targeted to what is most meaningful to different patients by providing different levels of support or additional treatment modules. The inclusion of telephone support may enable the clinician to assess patient preferences early in the program and adapt the course content to better suit their needs. Others have suggested that allowing the patient to choose the level of support they require would allow efficient allocation of clinician resources to where greater support is wanted, without affecting outcomes or adherence rates [50].

An important consideration related to patient adherence is the logical assumption that adherent patients will have better treatment outcomes than nonadherent patients, although the evidence has not always supported this assumption [51,52]. This study would suggest that additional telephone support does not have an effect on clinical outcomes; however, it did improve onboarding rates, thus increasing the number of participants who commenced the program. Web-based adherence is difficult to monitor because key therapy components are delivered on the web and practiced without the physical presence of a clinician [51]. For instance, a patient may practice the skills outlined in a program without logging in to the internet module. A more relevant way to measure adherence may be to look at intended use, defined as "the extent to which individuals *should* experience the content (of the intervention) to derive maximum benefit from the intervention, as defined or implied by its creators" [52]. This may explain why the patients who did not complete all modules in our study still showed significant improvements in their outcome measures. Further evaluation of the minimum amount of engagement with the program required to induce improvement may provide valuable insight. However, inherent difficulties lie in conducting such an analysis because those who stop engaging with treatments are also more likely to be lost to follow-up assessment.

As internet-delivered interventions evolve, the importance of various technical features in programs and their influence on adherence and outcomes need to be better understood. The mode of guidance, whether by email, text, phone, telehealth, program content, web design, multimedia format, web-based design features, or a hybrid of these features, may play a part in successful engagement of the individual [52]. Qualitative data would support this, with real-world Reboot Online service users anecdotally reporting a desire for more multimedia capability and web-based features. Program developers will need to consider these aspects of design that have an impact on usability and user experience during the development and ongoing evaluation of internet-delivered programs. Further evidence is needed to elucidate what effect each of these design components may have on user engagement and adherence and, in turn, whether design optimization strategies could be successfully used to improve treatment adherence and outcomes [20].

Studies investigating web-based CBT courses for pain conditions [24] report mixed results on whether adjunct guidance improves measures of pain catastrophizing, self-efficacy, and pain interference more than unguided internet interventions alone. Although our study showed significant improvements in measures of pain self-efficacy, kinesiophobia, pain interference, pain disability, and psychological distress over time, there was no significant difference in outcome measures between Reboot Online combined with telephone support and the usual Reboot Online intervention. This may indicate that adding telephone support may not add to the effect of the core Reboot Online program on clinical outcome measures.

The results presented here need to be considered within the context of a number of study limitations. Blinding of participants and the telephone support coach was not possible, introducing a possible risk of bias. Responder bias is another limitation noted because participants self-selected to participate in the trial and outcome data were lost from those participants who dropped out of the course. The power calculation was performed with reference to our primary outcome measure, which was adherence. The study was not powered formally for effectiveness on secondary outcomes and did not correct for multiple comparisons. Loss of follow-up data after treatment and at 3 months resulted in a smaller sample than estimated, which may have led to biased estimates or overestimates of treatment effects. Intention-to-treat analyses were conducted to enable inclusion of participants with suboptimal compliance into the data analysis, and linear mixed model analyses were selected, given that they remain robust in the presence of considerable missing data. Most of the baseline data were collected through patient-reported surveys, and thus the reliability of data are limited by the accuracy of participant recall and self-report. The subjective physical activity questionnaire may have been subject to notable reporting bias; an objective measure of physical activity (such as an activity-monitoring device) is needed in future evaluations' estimate of physical activity. Although outside the scope of this study, collection of more detailed data on direct intervention costs and opiate use could have facilitated more comprehensive evaluation of the benefits of the Reboot Online program. The study used 2 therapists (TG and JW) to conduct the telephone support, both senior clinicians in chronic pain and trained in the telephone coaching protocol. Although this allowed for consistent intervention during the study, it is unknown whether our findings would generalize to broader clinical settings or multiple clinicians. Finally, the study was conducted between September 2019 and November 2020. During this time period, there were significant and unprecedented local (drought and bushfires in Australia) and global events (COVID-19) that may have had an impact on the ability of participants to adhere and commit to the study as well as on measures of their well-being and mental and physical health.

Conclusions

Reboot Online offers an effective web-based intervention for chronic pain. Telephone support improves participants' registration, program commencement, and engagement in the early phase of the internet intervention; however, it did not seem to have an impact on overall course completion or efficacy. To maximize adherence to web-based interventions, clinician resources may be best used in the early phase of onboarding. Further research is warranted to gain better understanding of optimal guidance levels and models as well as program design components that will most effectively improve adherence to web-based interventions.

Acknowledgments

This project was supported by a grant from the St Vincent's Health Australia Inclusive Health Innovation Fund, St Vincent's Clinic Foundation, 2019. This Way Up is funded by the Australian Government Department of Health. JN is supported by an Australian Medical Research Future Fund Fellowship (MRFF1145382).

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 699 KB - jmir_v24i2e30880_app1.pdf]

References

1. Blyth F, March L, Brnabic A, Jorm L, Williamson M, Cousins M. Chronic pain in Australia: a prevalence study. Pain 2001;89(2):127-134. [doi: 10.1016/S0304-3959(00)00355-9]
2. González-Chica DA, Vanlint S, Hoon E, Stocks N. Epidemiology of arthritis, chronic back pain, gout, osteoporosis, spondyloarthropathies and rheumatoid arthritis among 1.5 million patients in Australian general practice: NPS MedicineWise MedicineInsight dataset. BMC Musculoskelet Disord 2018 Jan 18;19(1):20 [FREE Full text] [doi: 10.1186/s12891-018-1941-x] [Medline: 29347932]
3. Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. Eur J Pain 2006 May;10(4):287-333. [doi: 10.1016/j.ejpain.2005.06.009] [Medline: 16095934]
4. Vos T, Abajobir AA, Abate KH, Abbafati C, Abbas KM, Abd-Allah F, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: A systematic analysis for the global burden of disease study 2013. The Lancet 2015;386(9995):743-800. [doi: 10.1016/S0140-6736(17)32154-2]
5. Casey M, Smart K, Segurado R, Doody C. Multidisciplinary-Based Rehabilitation (MBR) compared with active physical interventions for pain and disability in adults with chronic pain: a systematic review and meta-analysis. Clin J Pain 2020 Nov;36(11):874-886. [doi: 10.1097/AJP.0000000000000871] [Medline: 32773436]
6. Garg S, Garg D, Turin T, Chowdhury M. Web-based interventions for chronic back pain: a systematic review. J Med Internet Res 2016 Jul 26;18(7):e139 [FREE Full text] [doi: 10.2196/jmir.4932] [Medline: 27460413]
7. Hruschak V, Flowers KM, Azizoddin DR, Jamison RN, Edwards RR, Schreiber KL. Cross-sectional study of psychosocial and pain-related variables among patients with chronic pain during a time of social distancing imposed by the coronavirus disease 2019 pandemic. Pain 2021 Feb 01;162(2):619-629 [FREE Full text] [doi: 10.1097/j.pain.0000000000002128] [Medline: 33230007]
8. Eccleston C, Fisher E, Craig L, Duggan G, Rosser B, Keogh E. Psychological therapies (Internet-delivered) for the management of chronic pain in adults. Cochrane Database Syst Rev 2012:CD010152. [doi: 10.1002/14651858.CD010152]
9. Bender J, Radhakrishnan A, Diorio C, Englesakis M, Jadad A. Can pain be managed through the internet? A systematic review of randomized controlled trials. Pain 2011 Aug;152(8):1740-1750. [doi: 10.1016/j.pain.2011.02.012] [Medline: 21565446]
10. Cuijpers P, van Straten A, Andersson G. Internet-administered cognitive behavior therapy for health problems: a systematic review. J Behav Med 2008 Apr;31(2):169-177 [FREE Full text] [doi: 10.1007/s10865-007-9144-1] [Medline: 18165893]
11. Vugts M, Joosen M, van der Geer JE, Zedlitz A, Vrijhoef H. The effectiveness of various computer-based interventions for patients with chronic pain or functional somatic syndromes: a systematic review and meta-analysis. PLoS One 2018;13(5):e0196467 [FREE Full text] [doi: 10.1371/journal.pone.0196467] [Medline: 29768436]
12. Murray E, Burns J, See T, Lai R, Nazareth I. Interactive Health Communication Applications for people with chronic disease. Cochrane Database Syst Rev 2005 Oct 19;4(4):CD004274. [doi: 10.1002/14651858.CD004274.pub4] [Medline: 16235356]

13. Wantland D, Portillo C, Holzemer W, Slaughter R, McGhee E. The effectiveness of web-based vs. non-web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40 [FREE Full text] [doi: 10.2196/jmir.6.4.e40] [Medline: 15631964]
14. Buhrman M, Fältenhag S, Ström L, Andersson G. Controlled trial of internet-based treatment with telephone support for chronic back pain. Pain 2004 Oct;111(3):368-377. [doi: 10.1016/j.pain.2004.07.021] [Medline: 15363881]
15. Schultz R, Smith J, Newby JM, Gardner T, Shiner CT, Andrews G, et al. Pilot trial of the reboot online program: an internet-delivered, multidisciplinary pain management program for chronic pain. Pain Res Manag 2018 Sep 05;2018:1-11 [FREE Full text] [doi: 10.1155/2018/9634727] [Medline: 30254712]
16. Smith J, Faux S, Gardner T, Hobbs M, James M, Joubert A, et al. Reboot online: a randomized controlled trial comparing an online multidisciplinary pain management program with usual care for chronic pain. Pain Med 2019 Dec 01;20(12):2385-2396. [doi: 10.1093/pm/pnz208] [Medline: 31498393]
17. Lim D, Newby J, Gardner T, Haskelberg H, Schultz R, Faux S, et al. Evaluating real-world adherence and effectiveness of the "Reboot Online" program for the management of chronic pain in routine care. Pain Med 2021 Aug 06;22(8):1784-1792. [doi: 10.1093/pm/pnaa458] [Medline: 33729536]
18. Andersson G. The Internet and CBT: A Clinical Guide. London, UK: CRC Press; 2014.
19. Buhrman M, Gordh T, Andersson G. Internet interventions for chronic pain including headache: a systematic review. Internet Interv 2016 May;4:17-34 [FREE Full text] [doi: 10.1016/j.invent.2015.12.001] [Medline: 30135787]
20. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on internet-based mental health interventions — A systematic review. Internet Intervent 2014 Oct;1(4):205-215. [doi: 10.1016/j.invent.2014.08.003]
21. Evans EH, Araújo-Soares V, Adamson A, Batterham AM, Brown H, Campbell M, et al. The NULevel trial of a scalable, technology-assisted weight loss maintenance intervention for obese adults after clinically significant weight loss: study protocol for a randomised controlled trial. Trials 2015 Sep 22;16(1):421 [FREE Full text] [doi: 10.1186/s13063-015-0931-7] [Medline: 26395774]
22. Barnett J, Harricharan M, Fletcher D, Gilchrist B, Coughlan J. myPace: an integrative health platform for supporting weight loss and maintenance behaviors. IEEE J Biomed Health Inform 2015 Jan;19(1):109-116. [doi: 10.1109/jbhi.2014.2366832]
23. Schäfer AG, Zalpour C, von Piekartz H, Hall TM, Paelke V. The efficacy of electronic health-supported home exercise interventions for patients with osteoarthritis of the knee: systematic review. J Med Internet Res 2018 Apr 26;20(4):e152 [FREE Full text] [doi: 10.2196/jmir.9465] [Medline: 29699963]
24. Lin J, Sander L, Paganini S, Schlicker S, Ebert D, Berking M, et al. Effectiveness and cost-effectiveness of a guided internet- and mobile-based depression intervention for individuals with chronic back pain: protocol of a multi-centre randomised controlled trial. BMJ Open 2017 Dec 28;7(12):e015226 [FREE Full text] [doi: 10.1136/bmjopen-2016-015226] [Medline: 29288172]
25. Stinson JN, McGrath PJ, Hodnett ED, Feldman BM, Duffy CM, Huber AM, et al. An internet-based self-management program with telephone support for adolescents with arthritis: a pilot randomized controlled trial. J Rheumatol 2010 Sep;37(9):1944-1952. [doi: 10.3899/jrheum.091327] [Medline: 20595280]
26. Learn practical tools to take care of your mental health. This way up. URL: https://thiswayup.org.au/ [accessed 2022-01-18]
27. Lecrubier Y, Sheehan D, Weiller E, Amorim P, Bonora I, Sheehan KH, et al. The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. Eur Psychiatr 2020 Apr 16;12(5):224-231. [doi: 10.1016/s0924-9338(97)83296-8]
28. Soer R, Reneman MF, Vroomen PC, Stegeman P, Coppes MH. Responsiveness and minimal clinically important change of the pain disability index in patients with chronic back pain. Spine 2012;37(8):711-715. [doi: 10.1097/brs.0b013e31822c8a7a]
29. International Physical Activity Questionnaire. The IPAQ. 2016. URL: https://sites.google.com/site/theipaq/ [accessed 2021-12-30]
30. Welcome to randomization.com!!!: where it's never the same thing twice. Randomization. 2020. URL: http://www.jerrydallal.com/random/randomize.htm [accessed 2022-01-18]
31. How REDCap is being used in response to COVID-19. REDCap. URL: https://www.project-redcap.org/ [accessed 2022-01-18]
32. The Chronic Pain Course. This way up. URL: https://thiswayup.org.au/courses/the-chronic-pain-course/ [accessed 2022-01-18]
33. Nicholas M. The pain self-efficacy questionnaire: taking pain into account. Eur J Pain 2007 Mar;11(2):153-163. [doi: 10.1016/j.ejpain.2005.12.008] [Medline: 16446108]
34. Chiarotto A, Vanti C, Cedraschi C, Ferrari S, de Lima E Sà Resende F, Ostelo RW, et al. Responsiveness and minimal important change of the pain self-efficacy questionnaire and short forms in patients with chronic low back pain. J Pain 2016 Jun;17(6):707-718. [doi: 10.1016/j.jpain.2016.02.012] [Medline: 26975193]
35. Miller R, Kori S, Todd D. The Tampa Scale: a measure of kinesiophobia. Clin J Pain 1991;7(1):51 [FREE Full text]
36. Monticone M, Ambrosini E, Rocca B, Foti C, Ferrante S. Responsiveness and minimal clinically important changes for the Tampa Scale of Kinesiophobia after lumbar fusion during cognitive behavioral rehabilitation. Eur J Phys Rehabil Med 2017 Jun;53(3):351-358. [doi: 10.23736/s1973-9087.16.04362-8]

37. Cleeland C, Ryan K. Pain assessment: global use of the brief pain inventory. Ann Acad Med Singapore 1994;23(2):129-138. [doi: 10.1097/01893697-199513010-00022]
38. Mease PJ, Spaeth M, Clauw DJ, Arnold LM, Bradley LA, Russell IJ, et al. Estimation of minimum clinically important difference for pain in fibromyalgia. Arthritis Care Res (Hoboken) 2011 Jun 31;63(6):821-826 [FREE Full text] [doi: 10.1002/acr.20449] [Medline: 21312349]
39. Chibnall J, Tait R. The Pain Disability Index: factor structure and normative data. Arch Phys Med Rehabil 1994 Oct;75(10):1082-1086. [doi: 10.1016/0003-9993(94)90082-5] [Medline: 7944912]
40. Craig CL, Marshall AL, Sjostrom M, Bauman AE, Booth ML, Ainsworth BE, et al. International Physical Activity Questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003;35(8):1381-1395. [doi: 10.1249/01.mss.0000078924.61453.fb]
41. Cheng HL. A simple, easy-to-use spreadsheet for automatic scoring of the International Physical Activity Questionnaire (IPAQ) Short Form. ResearchGate 2016. [doi: 10.13140/RG.2.2.21067.80165]
42. Furukawa TA, Kessler RC, Slade T, Andrews G. The performance of the K6 and K10 screening scales for psychological distress in the Australian National Survey of Mental Health and Well-Being. Psychol Med 2003 Mar 14;33(2):357-362. [doi: 10.1017/s0033291702006700] [Medline: 12622315]
43. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SL, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med 2002 Aug 26;32(6):959-976. [doi: 10.1017/s0033291702006074] [Medline: 12214795]
44. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
45. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
46. Cohen J. Statistical power analysis. Curr Dir Psychol Sci 2016 Jun 24;1(3):98-101. [doi: 10.1111/1467-8721.ep10768783]
47. Carlbring P, Maurin L, Törngren C, Linna E, Eriksson T, Sparthan E, et al. Individually-tailored, Internet-based treatment for anxiety disorders: a randomized controlled trial. Behav Res Ther 2011 Jan;49(1):18-24. [doi: 10.1016/j.brat.2010.10.002] [Medline: 21047620]
48. Spek V, Cuijpers P, Nyklícek I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. Psychol Med 2006 Nov 20;37(03):319. [doi: 10.1017/s0033291706008944]
49. Chien I, Enrique A, Palacios J, Regan T, Keegan D, Carter D, et al. A machine learning approach to understanding patterns of engagement with internet-delivered mental health interventions. JAMA Netw Open 2020 Jul 01;3(7):e2010791 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.10791] [Medline: 32678450]
50. Hadjistavropoulos H, Schneider L, Mehta S, Karin E, Dear B, Titov N. Preference trial of internet-delivered cognitive behaviour therapy comparing standard weekly versus optional weekly therapist support. J Anxiety Disord 2019 Apr;63:51-60 [FREE Full text] [doi: 10.1016/j.janxdis.2019.02.002] [Medline: 30844613]
51. Andersson G, Titov N. Advantages and limitations of internet-based interventions for common mental disorders. World Psychiatry 2014 Mar 04;13(1):4-11 [FREE Full text] [doi: 10.1002/wps.20083] [Medline: 24497236]
52. Sieverink F, Kelders SM, van Gemert-Pijnen JE. Clarifying the concept of adherence to eHealth technology: systematic review on when usage becomes adherence. J Med Internet Res 2017 Dec 06;19(12):e402 [FREE Full text] [doi: 10.2196/jmir.8578] [Medline: 29212630]

Abbreviations

CBT: cognitive behavioral therapy
CONSORT: Consolidated Standards of Reporting Trials
MCID: minimally clinically important difference
MDD: major depressive disorder
REDCap: Research Electronic Data Capture
SMART: specific, measurable, achievable, relevant, and timely

Edited by A Mavragani; submitted 07.06.21; peer-reviewed by A Martin-Pintado Zugasti, S Bhattacharjee, A Kebede; comments to author 01.08.21; revised version received 24.08.21; accepted 07.12.21; published 03.02.22.

Please cite as:
Gardner T, Schultz R, Haskelberg H, Newby JM, Wheatley J, Millard M, Faux SG, Shiner CT
The Effect of Adjunct Telephone Support on Adherence and Outcomes of the Reboot Online Pain Management Program: Randomized Controlled Trial
J Med Internet Res 2022;24(2):e30880
URL: https://www.jmir.org/2022/2/e30880
doi:10.2196/30880
PMID:35113021

Original Paper

Effectiveness of Unguided Internet-Based Cognitive Behavioral Therapy and the Three Good Things Exercise for Insomnia: 3-Arm Randomized Controlled Trial

Daisuke Sato[1,2], PhD; Yoichi Sekizawa[3], MA; Chihiro Sutoh[1], MD, PhD; Yoshiyuki Hirano[4,5], PhD; Sho Okawa[4], PhD; Motohisa Hirose[1,4], PhD; Ryo Takemura[6], PhD; Eiji Shimizu[1,2,4,5], MD, PhD

[1]Department of Cognitive Behavioral Physiology, Graduate School of Medicine, Chiba University, Chiba, Japan

[2]Cognitive Behavioral Therapy Center, Chiba University Hospital, Chiba, Japan

[3]Research Institute of Economy, Trade and Industry, Tokyo, Japan

[4]Research Center for Child Mental Development, Chiba University, Chiba, Japan

[5]United Graduate School of Child Development, Osaka University, Kanazawa University, Hamamatsu University School of Medicine, Chiba University and University of Fukui, Suita, Japan

[6]Department of Preventive Medicine and Public Health, Keio University School of Medicine, Tokyo, Japan

Corresponding Author:
Daisuke Sato, PhD
Department of Cognitive Behavioral Physiology, Graduate School of Medicine
Chiba University
1-8-1 Inohana
Chuo-ku, Chiba-shi
Chiba, 260-8670
Japan
Phone: 81 043 226 2027
Fax: 81 043 226 2028
Email: daisuke-sato@umin.ac.jp

Abstract

Background: The treatment of insomnia with sleep medication causes problems such as long-term use, dependence, and significant economic losses, including medical expenses. Evidence-based lifestyle guidance is required to improve insomnia symptoms not only in person but also in easy-to-use web-based formats.

Objective: This study aims to clarify whether unguided internet-based cognitive behavioral therapy (ICBT) or the Three Good Things (TGT) exercise, both administered as self-help internet interventions without email support, could improve insomnia symptoms compared with a waiting list control (WLC) group.

Methods: A 4-week program was implemented, and participants were randomly allocated to 1 of the 3 groups. The primary outcome measure was the Pittsburgh Sleep Questionnaire (PSQI) score at 4 weeks compared with baseline.

Results: Of the 21,394 individuals invited to participate, 312 (1.46%) met the eligibility criteria and were randomly assigned to 1 of the 3 groups. Of these 312 individuals, 270 (86.5%; ICBT 79/270, 29.3%; TGT 88/270, 32.6%; and WLC 103/270, 38.1%) completed a postintervention survey at 4 and 8 weeks. The adjusted mean changes of the primary outcome measure (PSQI) in the ICBT (−1.56, 95% CI −2.52 to −0.59; P<.001) and TGT (−1.15, 95% CI −2.08 to −0.23; P=.002) groups at 4 weeks from baseline showed a significant improvement compared with the WLC group. The adjusted mean changes in the secondary outcome measures of sleep onset latency, total sleep time, Athens Insomnia Scale score, and Patient Health Questionnaire-9 score at 4 weeks from baseline, as well as in the PSQI at 8 weeks from baseline, showed significant improvement for ICBT. Moreover, total sleep time, Athens Insomnia Scale, and Patient Health Questionnaire-9 scores at 4 weeks from baseline showed a significant improvement in the TGT group compared with the WLC group.

Conclusions: A total of 4 weeks of unguided ICBT and TGT exercises improved insomnia.

Trial Registration: University Hospital Medical Information Network Clinical Trial Registry UMIN000034927; https://center6.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000039814

(J Med Internet Res 2022;24(2):e28747) doi:10.2196/28747

KEYWORDS

insomnia; internet-based treatment; cognitive behavioral therapy; positive psychology; randomized controlled trial; mobile phone

Introduction

Background

Insomnia describes the inability to sleep, including difficulty falling asleep, difficulty staying asleep, difficulty sleeping, and early waking, despite an attempt to sleep at the right time and in the right environment. This condition reduces an individual's quality of life by causing various daytime dysfunctions, including sleepiness during the day [1]. Insomnia can be caused by a variety of factors, including changes in social life such as decreased interpersonal interaction and socializing, as well as mental and physical illnesses, decreased physical function, and physiological changes in the brain. Insomnia impairs daytime activity and triggers additional illnesses such as depression, lifestyle-related diseases, and cancer. Insomnia is a common complaint in adults, with 13.5% to 20.6% of the general population experiencing daytime sleepiness >3 times per week [2]. It has been reported that 14.9% of people in Japan have profound daytime sleepiness [3]. Another study found that 1 in 5 Japanese adults had sleep problems [4]. In Japan, the economic losses caused by sleep deprivation and insomnia have been estimated to be 15 trillion yen (US $8.67 billion) [5,6].

From the perspective of optimizing medical costs associated with lifestyle-related diseases, such as diabetes, hypertension, and hyperlipidemia, lifestyle guidance is increasingly being introduced as a complement to the use of medication. However, the evidence base to support the introduction of lifestyle guidance is lacking. The provision of face-to-face advice paired with an insomnia improvement program delivered as easy-to-use internet-based life guidance will likely reduce medical expenses.

Insomnia practice guidelines in the United States and Australia recommend cognitive behavioral therapy (CBT) as the first choice for the treatment of insomnia. In Japan, there are very few medical institutions that can provide CBT, and there are no barriers to prescribing psychotropic drugs outside of psychiatry and psychosomatic medicine. However, pharmacotherapy is associated with side effects such as dependence, tolerance, anterograde amnesia, muscle relaxant effects, carryover effects, and rebound insomnia. Furthermore, it increases the risk of double prescription and overdose and requires long-term treatment. However, the success of pharmacotherapy is limited. According to the Japanese Society of Sleep Studies guidelines, it is preferable to prioritize CBT over pharmacotherapy for insomnia.

CBT for insomnia is a treatment that focuses on anxiety and biased thinking about sleep and improves insomnia by reviewing an individual's lifestyle, existing anxiety and tension, and any thinking related to the maintenance of insomnia. Sleep hygiene instructions can provide important knowledge about sleep, including how to adjust living conditions so that good-quality sleep can be obtained, and practice measures for maintaining good-quality sleep throughout life. This study used internet-based CBT (ICBT) without email support to provide sleep hygiene guidance as a trial treatment. ICBT was paired with a self-help intervention as a form of noninvasive self-medication.

Positive psychology focuses on positive psychological traits such as happiness, optimism, and satisfaction with life and was developed by Seligman et al [7] of the University of Pennsylvania based on the reflection that traditional psychology was excessively focused on normalizing negative traits such as mental illness. Positive psychology is based on the idea of nurturing [7]. A typical method of positive psychology comprises the exercise of writing 3 good things every day before going to bed every night [7]. It is a simple diary-like exercise that involves listing 3 good things that happened that day and providing a written explanation of why those things happened. In this study, this exercise was called the Three Good Things (TGT) exercise.

ICBT has gained increasing attention as a possible treatment for improving insomnia, and many studies on ICBT have been conducted. ICBT provides web-based CBT and sleep hygiene guidance, simulating an ordinary CBT session (counseling) that is typically conducted by the patient and therapist face-to-face. With the approval of the Chiba University School of Medicine Hospital clinical trials ethics review board (approval number G27040), the combination of guided ICBT for insomnia and routine care (ie, usual care [UC]) has been shown to significantly reduce the Pittsburgh Sleep Quality Index (PSQI) score [8,9]. In a randomized controlled trial (RCT) of 23 patients with insomnia whose symptoms persisted even after taking sleep medications such as benzodiazepines, the adjusted mean change in total PSQI score from baseline to 6 weeks was −6.11 in the ICBT+UC group (n=11), which was significantly better than the 0.40 change seen in those who underwent UC alone (n=12; $P<.001$). In addition, a significant improvement was seen for the adjusted mean changes in PSQI score; sleep onset latency (SOL); sleep efficiency (SE); number of awakenings; and mean change in depression at 3, 6, and 12 weeks from baseline in the ICBT+UC group. No adverse events were reported.

Seligman et al [7] previously evaluated the effects of the TGT exercise in a RCT. According to that study, the group that performed this exercise daily for 1 week not only increased self-reported levels of happiness but also reported decreased depression compared with the control group, and the effect persisted after 6 months. To the best of our knowledge, no studies have examined whether the TGT exercise can reduce insomnia. However, recent reviews have shown that there is a correlation between positive emotions and sleep [10]. TGT may improve sleep by improving positive emotions.

Objectives

This study aims to determine the effectiveness of an unguided, web-based self-help intervention in reducing insomnia. Using an RCT design, a nonclinical population of adults with insomnia

was allocated to one of three groups: (1) the ICBT group, (2) the TGT group, and (3) the waitlist group. It is hypothesized that unguided ICBT and TGT would result in higher quality sleep in the treatment groups than in the waitlist group.

Methods

Trial Design

We report this RCT trial in accordance with the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) version 1.6 checklist (Multimedia Appendix 1).

This study comprised an exploratory, parallel-group (3 groups), randomized, open-label, controlled study, incorporating a nonintervention group (waiting list control [WLC]). The registration of the participants was started in February 2019.

Participants and Recruitment

An email was sent to the registered monitors owned by the internet research company that commissioned the research, and a preliminary survey was conducted on the internet in relation to the inclusion and exclusion criteria given in the following sections, after which informed consent was obtained.

Potential participants were required to meet all the following inclusion criteria at the time of the preliminary survey: (1) aged 20 to 70 years (regardless of sex); (2) patients with mild or severe sleep disorders, who scored ≥6 on the Athens Insomnia Scale (AIS) and ≥6 on the PSQI; (3) sleep disorder occurring at least three nights per week and lasting for at least 3 months; (4) access to internet use with PCs, smartphones, and tablets; and (5) understood the explanations in Japanese and freely provided web-based consent.

Exclusion criteria of the study included (1) those with moderate or severe anxiety symptoms, scoring ≥10 on the Generalized Anxiety Disorder-7 (GAD-7) questionnaire; (2) those with moderate or severe depression symptoms, scoring ≥10 points on the Patient Health Questionnaire-9 (PHQ-9); (3) the confirmed presence of alcohol or drug dependence, diagnosed by a medical institution within the past year, excluding tobacco smokers; (4) suicidal thoughts, with a score of ≥2 in the Q9 part of the PHQ-9; (5) the confirmed diagnosis of cerebral organic diseases such as sleep apnea syndrome, restless leg syndrome, epilepsy, dementia, or cerebrovascular disease; (6) the confirmed diagnosis of a severe progressive physical illness, such as cancer or heart failure; (7) the use of pharmacotherapy for mental disorders, including insomnia symptoms; and (8) night shift work between 8 PM and 2 AM.

Individuals who met the above conditions and provided consent were enrolled as study participants. This study excluded patients commonly treated with pharmacotherapy for insomnia.

Case Registration and Allocation Methods

Case registration was performed via the trial website. Among the registered members (candidate participants) of monitors owned by the internet research company, those who met the conditions were automatically assigned participant identification codes when they accessed the servers. Members voluntarily underwent a preliminary survey after explanation and agreement, following the distribution of the recruitment materials. In the first preliminary survey, AIS, PSQI, GAD-7, and PHQ-9 were administered to select and screen participants for inclusion. Only those who met the inclusion criteria and did not meet any of the exclusion criteria were able to proceed to the next preliminary survey. The second preliminary survey was conducted 2 weeks after the first preliminary survey and contained the same content as the first survey. Only those who met the inclusion criteria and did not meet any of the exclusion criteria proceeded to take part in the study. Responses from the second preliminary survey were used in the analysis as baseline values before the intervention. After the participants were randomly assigned, the method of accessing the intervention program assigned to each participant was provided by email.

Randomization

Allocation modifiers included baseline PSQI scores (≥12 and ≤11) and gender (male and female).

Intervention

Intervention Schedule and Methods

Participants were randomly assigned to 1 of the 3 groups. The study period was 4 weeks; thus, the ICBT group underwent an unguided ICBT program for 4 weeks, the TGT group underwent a TGT exercise program for 4 weeks, and the nonintervention group (WLC) waited for 4 weeks without an intervention. They then received access to their assigned unguided ICBT program and the TGT exercise program for 4 weeks, 7 times a week for 15 to 20 minutes each time.

The intervention programs comprised the following two types: unguided ICBT program and TGT program.

Unguided ICBT Program

We encouraged participants to take part in the unguided ICBT program for 4 weeks. Participants accessed the ICBT site and performed the following tasks autonomously:

1. Week 1 session: keeping a sleep diary (understanding the significance of the sleep diary and how to write it)
2. Week 2 session: changing behavior (improving behavioral habits that maintain insomnia, using stimulus control)
3. Week 3 session: capturing thoughts (reconstructing cognition related to sleep using the column method)
4. Week 4 session: changing sleep time (adjusting sleep schedule using the sleep restriction method)

TGT Program

Participants were encouraged to participate in the TGT exercise for 4 weeks. The participants who accessed the site received the prompt by email and autonomously filled out the 3 good things on that day. An explanation of the methods of the TGT exercise is provided elsewhere by Seligman et al [7] and Seligman [11].

Measures

Evaluation Items

The primary outcome was the change in PSQI scores from baseline to the postintervention survey at 4 weeks. The

secondary outcome was the change in PSQI scores from baseline to the postintervention survey at 8 weeks and changes in SOL, total sleep time (TST), SE, AIS, GAD-7, PHQ-9, and Center for Epidemiologic Studies–Depression Scale (CES-D) positive items only from baseline to the first and second postintervention surveys at 4 and 8 weeks. Web-based assessments were also administered.

PSQI Evaluation

The PSQI is a self-administered questionnaire for evaluating sleep and sleep quality [12]. Eighteen questions evaluated seven factors: sleep quality, sleep onset time, sleep time, SE, difficulty sleeping, use of sleep medication, and drowsiness during the day that hinders daily life. A total of 7 factors (0-3 points) were summed to calculate the total score (0-21 points). The higher the score, the greater the sleep impairment. A Japanese version has also been developed, with the cutoff set at 6 points [13].

AIS Evaluation

The AIS is a self-administered questionnaire for assessing the severity of insomnia [14]. It comprises eight items: sleep, midnight awakening, early morning awakening, TST, overall sleep quality, daytime satisfaction, physical and mental daytime activity, and daytime sleepiness. A total of 8 items (0-3 points) were summed to calculate the total score (0-24 points). A Japanese version has also been developed, with the cutoff set at 6 points [15].

GAD-7 Evaluation

The GAD-7 is a self-administered questionnaire for evaluating generalized anxiety disorders [16]. It comprises 7 items, and covers the 2 weeks immediately before the administration of the test. The scoring system is as follows: *No at all: 0 points*, *Several days: 1 point*, *More than half: 2 points*, and *Almost every day: 3 points*. The total score is calculated from 0 to 21 points. A Japanese version has also been developed, with 0 to 4 points suggesting no generalized anxiety disorder, 5 to 9 points representing mild generalized anxiety disorder, 10 to 14 representing moderate generalized anxiety disorder, and 15 to 21 indicating severe generalized anxiety disorder [17].

PHQ-9 Evaluation

The PHQ-9 is a self-administered questionnaire for evaluating major depressive disorder [18]. It comprises 9 items and covers 2 weeks immediately before administration of the test. The scoring system is as follows: *No at all: 0 points*, *Several days: 1 point*, *Half or more: 2 points*, and *Almost every day: 3 points*. The total score is calculated from 0 to 27 points. A Japanese version has also been developed, with 0 to 4 points suggesting that the individual is not depressed, 5 to 9 points suggesting mild depression, 10 to 14 points suggesting moderate depression, 15 to 19 points suggesting moderate to severe depression, and 20 to 27 points suggesting severe depression [19].

CES-D Evaluation

The CES-D is a self-administered questionnaire for assessing depression [20]. It comprises 20 items about mood and physical condition covering the past week using a 4-point scale ranging from 0 (not at all) to 3 (over 5 days). In this study, 16 negative items were not used, and only positive items were used. The positive items comprised four statements: “I think I have the same ability as other people,” “I can think positively about the future,” “I can enjoy my life without complaints,” and “I enjoy every day.” The total score is calculated from 0 to 12 points. A Japanese version has also been developed, with a higher total score indicating a higher degree of positive emotions [21].

Statistical Analysis

Overview

All participants who enrolled in this study, responded to the intervention program at least once after randomization, and had efficacy data were the most significant population for full analysis set (FAS). However, participants for whom baseline data could not be obtained and those who violated the dominant protocol (eg, violating the inclusion criteria or exclusion criteria and incorrect assignment to the intervention program) were excluded.

Participants selected from the FAS and those meeting the following criteria were selected as the target population per-protocol set (PPS) that complied with the study protocol: (1) completed at least 75% of the intervention program, (2) had available measurements of critical variables, and (3) had no major test plan violation such as selection criteria violation, exclusion criteria violation, or incorrect assignment to an intervention program.

The sample size was based on a previous study by van Straten et al [22], which indicated that the estimated difference in changes of PSQI scores from baseline was approximately 3.5. Assuming a group difference of 3.5 (SD 10.0) points, 66 participants per group provided 80% power to detect a difference in PSQI scores among the WLC, ICBT, and TGT groups at a 5% significance level.

Analysis of Participant Background

The distribution of participant background data and summary statistics for each analysis population were calculated for each group. For nominal variables, the frequency and proportion of categories have been shown for each group. For continuous variables, summary statistics (number of cases, mean, SD, minimum, median, and maximum) were calculated for each group. Pearson chi-square test was used for nominal variables, except when ≥20% of the cells had an expected frequency <5 (when Fisher exact test was used). 2-tailed Student *t* test, or Wilcoxon rank-sum test, were also used. The significance level was set at 5% for both sides.

Analysis of Effectiveness

The primary outcome of the efficacy of the interventions was the PSQI score, an index for improving insomnia. The primary purpose of this study was to examine the superiority of the ICBT and TGT groups in improving the insomnia status of nonclinical cases of insomnia compared with the WLC group. We estimated the difference in PSQI score change between the test and control groups and the 95% 2-sided CI. To test the null hypothesis that the changes in PSQI scores in both groups were equal in the primary analysis, covariance analysis was performed using the assignment adjustment factor as a covariate. The allocation adjustment factors were the PSQI score at the time of

registration and gender. The significance level of the test was set at 5% (2-sided). Adjustment of multiplicity was performed using the Dunnett method, as a pairwise comparison of the 2 groups was performed for the control group.

The secondary evaluation items of effectiveness were analyzed to supplement the primary analysis results. No adjustment for multiplicity was made in the analysis of the secondary efficacy outcomes. The significance level of the hypothesis test was set at 5% (2-sided), and the two-sided 95% CI was calculated.

Statement of Ethics

This study was approved by the clinical trials ethics review board of the Chiba University Hospital (registration number G30022) and was registered as a clinical trial (UMIN000034927). A document explaining consent was presented to the participants on a webpage, accompanied by a verbal explanation as part of a video animation by the principal investigator. After viewing these materials, individuals who freely agreed to participate were recruited into the study.

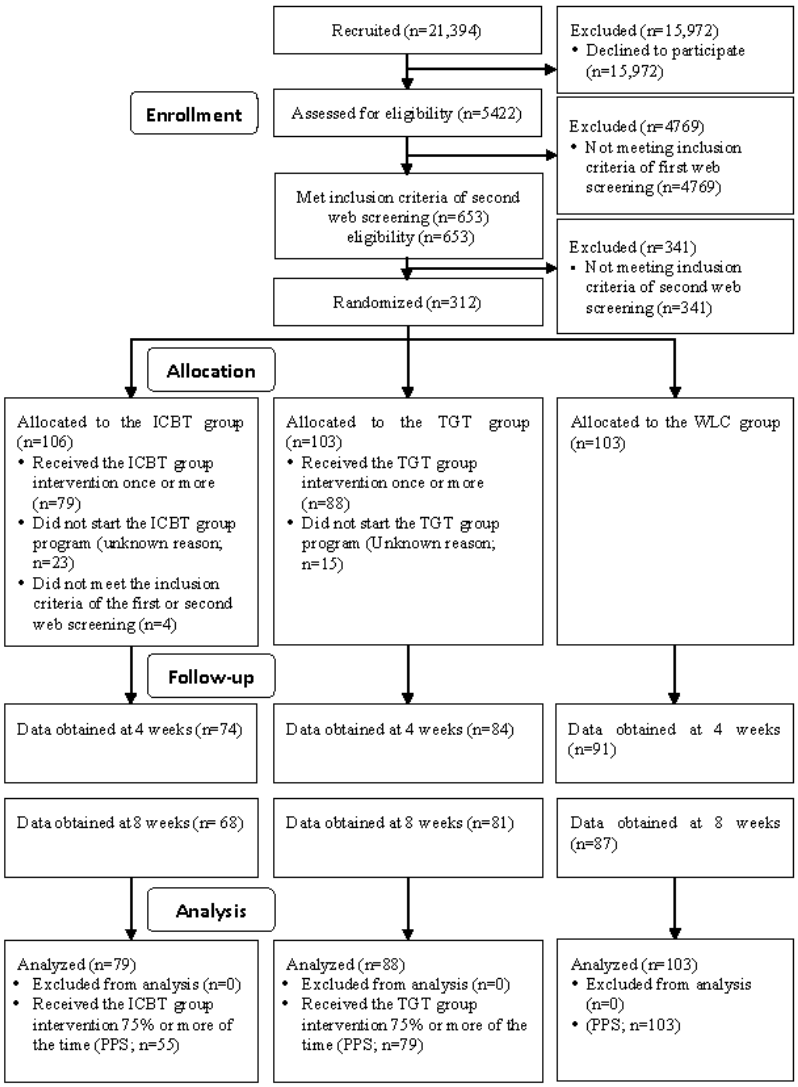
The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Results

Trial Flow

Figure 1 shows the research flow based on the CONSORT (Consolidated Standards of Reporting Trials) guidelines. Of the 21,394 individuals contacted through the internet research company, consent to participate in the study was obtained from 312 (1.46%) individuals who met the eligibility criteria and were randomly assigned to the ICBT, TGT, or WLC group. Of the 106 people assigned to ICBT, 23 (21.7%) did not start the ICBT program at least once, and 4 (3.8%) who performed the ICBT program at least once were classified as trial deviations, leaving 79 (74.5%) individuals with a FAS.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram of participant flow throughout the study. ICBT: internet-based cognitive behavioral therapy; PPS: per-protocol set; TGT: Three Good Things; WLC: waiting list control.



For details of the deviation related to PSQI, there was an error in the formula set by the internet survey company that conducted the web questionnaire, and the correct score was ≤5; however, it was judged to be ≥6 by incorrect calculation. Therefore, of

the 312 individuals, 4 (1.3%) individuals who had initially been selected as participants and who should not have been selected underwent the ICBT program; these individuals were reported as deviations to the Chiba University Hospital clinical trials ethics review committee and accepted as such.

Separately, there was an error in the formula used by the internet survey company that conducted the web questionnaire. In the error, the correct PSQI total score was ≥6; however, it was judged incorrectly to be ≤5. The correct PHQ-9 total score should have been calculated from the 1st to 9th items (the 10th item should have been excluded); however, the internet survey company that conducted the web questionnaire mistakenly calculated from the 1st to the 10th item. In addition, the correct GAD-7 total score should have been calculated from the first to seventh items (the eighth item was not added); however, the internet survey company that conducted the web questionnaire mistakenly calculated from the first to eighth items. For these reasons, of the 21,394 individuals, 83 (0.39%) were not selected in the first screening, and 25 (0.12%) were not selected in the second screening.

Of the 106 participants assigned to the ICBT group, 74 (69.8%) provided data at 4 weeks, 68 (64.2%) provided data at 8 weeks, and 55 (51.9%) had an attendance of ≥75% during the ICBT program; thus, the PPS was 55. Of the 103 participants assigned to the TGT group, 15 (14.6%) were excluded who did not start the TGT program at least once; thus, the FAS was 88. In the TGT group, of the 103 participants, 84 (79.2%) provided data at 4 weeks, and 81 (78.6%) provided data at 8 weeks. Approximately 76.7% (79/103) of participants had an attendance of ≥75% during the TGT program; thus, the PPS was 79.

Of the 312 participants, 103 (33%) were assigned to the WLC group; thus, the FAS score for this group was 103. Data were obtained successfully for 88.3% (91/103) of these participants at 4 weeks and for 84.5% (87/103) of participants at 8 weeks; thus, the PPS was 103. The registration of participants started in February 2019, and the follow-up ended in May 2019.

There were no adverse events. When adverse events occurred for participants, we asked them to report the adverse events by email. However, there were no reports.

Participant Characteristics

Table 1 shows the characteristics of the participants at baseline. Female participants comprised 41.5% (112/270; ICBT group 32/79, 41%; TGT group 39/88, 44%; and WLC group 41/103, 39.8%) of the study population, with a mean age of 50.4 (SD 10.8; ICBT group: 49.8, SD 11.1; TGT group: 50.5, SD 11.0; and WLC group: 51.0, SD 10.4) years. There were no significant differences in sex, age, marital status, educational history, or employment status among the 3 groups.

Table 2 presents the baseline data for the primary and secondary outcome measures. There were no significant differences among the 3 groups for the PSQI, AIS, GAD-7, or PHQ-9 scores.

Table 1. Participant characteristics (N=270).

Characteristics	ICBT[a] group (n=79)	TGT[b] group (n=88)	WLC[c] group (n=103)	*P* value
Sex , n (%)				.81
Female	32 (40.5)	39 (44.3)	41 (39.8)	
Male	47 (59.5)	49 (55.7)	62 (60.2)	
Age (years), mean (SD)	49.8 (11.1)	50.5 (11.0)	51.0 (10.4)	.78
Marriage, n (%)				.26
Single	19 (24.1)	18 (20.5)	25 (40.3)	
Married	58 (73.4)	60 (68.2)	72 (37.9)	
Divorce	2 (2.5)	10 (11.4)	6 (33.3)	
Education, mean (SD)	15.4 (1.9)	15.0 (2.0)	15.4 (2.1)	.33
Working, n (%)				.71
Full time	52 (65.8)	51 (57.9)	66 (39.1)	
Part-time	11 (13.9)	14 (15.9)	18 (41.9)	
Unemployed	16 (20.3)	23 (26.1)	19 (32.8)	

[a]ICBT: internet-based cognitive behavioral therapy.

[b]TGT: Three Good Things.

[c]WLC: waiting list control.

Table 2. Baseline data of primary outcome and secondary outcome (N=270).

Outcomes	Value, mean (SD)			*P* value
	ICBT[a] group (n=79)	TGT[b] group (n=88)	WLC[c] group (n=103)	
Primary outcome: insomnia symptoms (PSQI[d])	9.8 (2.4)	9.8 (2.0)	9.8 (2.3)	.97
Secondary outcome : sleep				
AIS[e]	10.2 (2.9)	10.0 (2.7)	10.2 (2.9)	.77
SOL[f] (minutes)	41.6 (27.9)	47.8 (40.6)	41.2 (35.0)	.66
TST[g] (hours)	318.6 (63.7)	323.8 (52.1)	319.9 (61.7)	.17
Sleep efficiency (%)	82.1 (15.2)	85.3 (10.9)	84.3 (14.7)	.55
Secondary outcome : mood				
GAD-7[h]	3.0 (2.5)	3.1 (2.3)	3.2 (2.5)	.92
PHQ-9[i]	4.5 (2.2)	4.6 (2.2)	4.4 (2.2)	.74

[a]ICBT: internet-based cognitive behavioral therapy.

[b]TGT: Three Good Things.

[c]WLC: waiting list control.

[d]PSQI: Pittsburgh Sleep Quality Index.

[e]AIS: Athens Insomnia Scale.

[f]SOL: sleep onset latency.

[g]TST: total sleep time.

[h]GAD-7: Generalized Anxiety Disorder-7.

[i]PHQ-9: Patient Health Questionnaire-9.

Intervention Effects

Primary Outcome

Figure 2 and Table 3 show the changes in the primary outcome score (PSQI) in each group, and Table 4 shows the comparison of the changes in the primary outcome PSQI between the intervention and control groups. At week 4, the adjusted mean change from baseline in the ICBT group compared with the WLC group was −1.56 (95% CI −2.52 to −0.59; *P*<.001), and the adjusted mean change from baseline in the TGT group compared with the WLC group was −1.15 (95% CI −2.08 to −0.23; *P*=.002), indicating that both the ICBT and TGT groups had a significant reduction in their PSQI scores compared with the WLC group.

Figure 2. Means and SDs (raw data) for the primary outcome and improvement in the Pittsburgh Sleep Quality Index (PSQI) score. ICBT: internet-based cognitive behavioral therapy; TGT: Three Good Things; WLC: waiting list control.

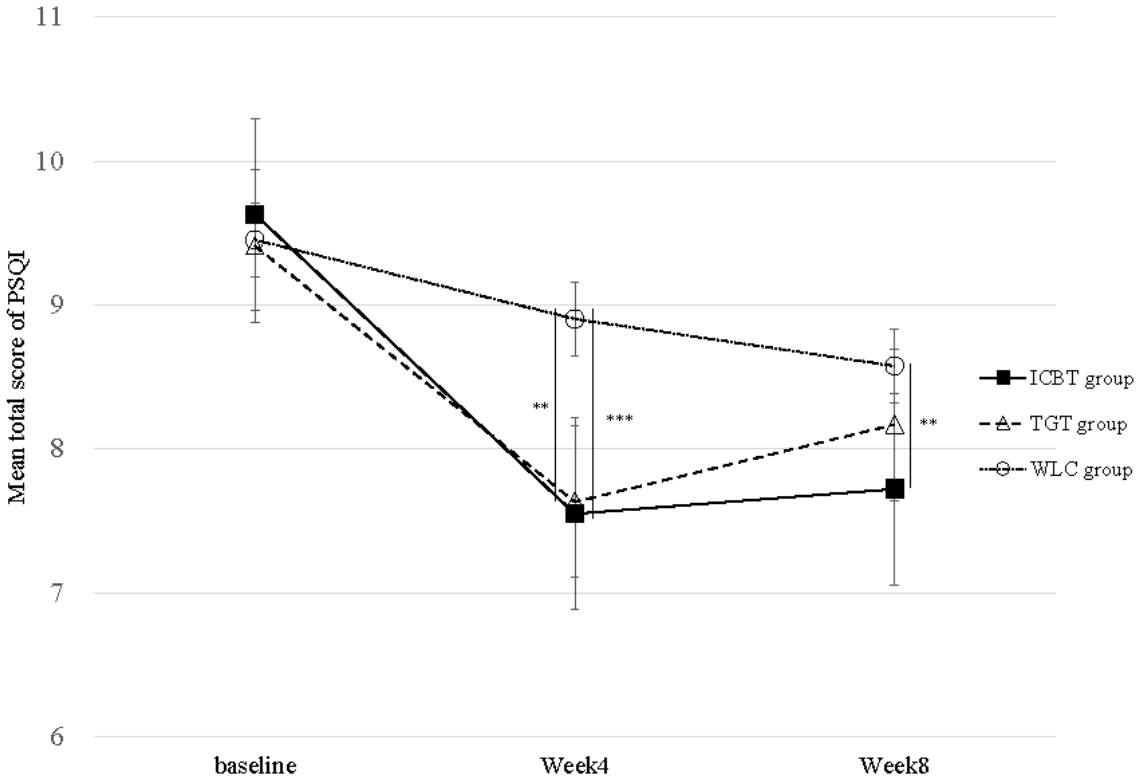


Table 3. Changes in primary outcome (PSQI) in the 3 groups.

Assigned group (week 4)	Estimation (SE; 95% CI)	*P* value
ICBT[a]	−2.19 (0.29; −2.76 to −1.63)	<.001
TGT[b]	−1.79 (0.26; −2.31 to –1.27)	<.001
WLC[c]	−0.64 (0.25; −1.14 to −0.14)	.01

[a]ICBT: internet-based cognitive behavioral therapy.

[b]TGT: Three Good Things.

[c]WLC: waiting list control.

Table 4. Changes in the secondary outcome among the 3 groups.

Secondary outcomes and assigned group	Estimation (SE)	*P* value	95% CI
PSQI[a]			
Week 8			
ICBT[b]	−2.09 (0.29)	<.001	−2.66 to −1.51
TGT[c]	−1.34 (0.26)	<.001	−1.86 to −0.82
WLC[d]	−0.99 (0.25)	<.001	−1.49 to −0.50
SOL[e]			
Week 4			
ICBT	−19.11 (2.55)	<.001	−24.13 to −14.08
TGT	−13.26 (2.38)	<.001	−17.95 to −8.56
WLC	−7.38 (2.29)	.001	−11.88 to −2.87
Week 8			
ICBT	−15.13 (2.64)	<.001	−20.33 to −9.94
TGT	−12.39 (2.40)	<.001	−17.11 to −7.67
WLC	−7.40 (2.26)	.001	−11.86 to −2.94
TST[f]			
Week 4			
ICBT	37.40 (5.27)	<.001	27.01 to 47.78
TGT	25.02 (4.89)	<.001	15.39 to 34.66
WLC	4.32 (4.72)	.36	−4.98 to 13.63
Week 8			
ICBT	31.73 (5.43)	<.001	21.03 to 42.43
TGT	14.25 (4.91)	.004	4.57 to 23.93
WLC	14.23 (4.68)	.003	5.00 to 23.45
Sleep efficiency			
Week 4			
ICBT	5.85 (1.37)	<.001	3.15 to 8.55
TGT	6.37 (1.28)	<.001	3.86 to 8.89
WLC	2.90 (1.22)	.02	0.50 to 5.31
Week 8			
ICBT	6.73 (1.42)	<.001	3.93 to 9.52
TGT	3.77 (1.30)	.004	1.19 to 6.34
WLC	2.19 (1.21)	.07	−0.19 to 4.57
AIS[g]			
Week 4			
ICBT	−2.11 (0.37)	<.001	−2.83 to −1.39
TGT	−2.75 (0.34)	<.001	−3.43 to −2.08
WLC	−0.49 (0.32)	<.001	−1.12 to 0.15
Week 8			
ICBT	−1.82 (0.38)	<.001	−2.57 to −1.08
TGT	−1.62 (0.34)	<.001	−2.29 to −0.95
WLC	−1.21 (0.32)	<.001	−1.84 to −0.57

Secondary outcomes and assigned group	Estimation (SE)	*P* value	95% CI
GAD-7[h]			
Week 4			
ICBT	0.17 (0.30)	.58	−0.43 to 0.77
TGT	−0.44 (0.28)	.12	−1.00 to 0.12
WLC	0.39 (0.27)	.15	−0.14 to 0.91
Week 8			
ICBT	0.07 (0.31)	.83	−0.55 to 0.69
TGT	0.11 (0.28)	.70	−0.45 to 0.67
WLC	0.07 (0.27)	.79	−0.45 to 0.60
PHQ-9[i]			
Week 4			
ICBT	−0.32 (0.35)	.37	−1.02 to 0.38
TGT	−0.40 (0.33)	.23	−1.05 to 0.25
WLC	1.38 (0.31)	<.001	0.77 to 2.00
Week 8			
ICBT	−0.27 (0.37)	.46	−0.99 to 0.45
TGT	−0.03 (0.33)	.93	−0.69 to 0.63
WLC	0.38 (0.31)	.22	−0.24 to 1.00
CES-D[j]			
Week 4			
ICBT	0.40 (0.36)	.27	−0.30 to 1.10
TGT	0.54 (0.33)	.10	−0.11 to 1.20
WLC	0.53 (0.31)	.09	−0.09 to 1.14
Week 8			
ICBT	0.47 (0.37)	.20	−0.25 to 1.19
TGT	0.52 (0.33)	.12	−0.13 to 1.18
WLC	−0.51 (0.31)	.11	−1.12 to 0.11

[a]PSQI: Pittsburgh Sleep Quality Index.

[b]ICBT: internet-based cognitive behavioral therapy.

[c]TGT: Three Good Things.

[d]WLC: waiting list control.

[e]SOL: sleep onset latency.

[f]TST: total sleep time.

[g]AIS: Athens Insomnia Scale.

[h]GAD-7: Generalized Anxiety Disorder-7.

[i]PHQ-9: Patient Health Questionnaire-9.

[j]CES-D: Center for Epidemiologic Studies–Depression Scale.

Secondary Outcomes

Table 4 shows the changes in the secondary evaluation items for each group. Table 5 shows a comparison of the changes in secondary outcomes between the intervention and control groups.

Table 5. Comparison of changes in the secondary outcome between intervention groups and control group.

Secondary outcomes and assigned group	Control group	Estimation (SE)	Adjusted *P* value[a]	95% CI (adjusted)
PSQI[b]				
Week 8				
ICBT[c]	WLC[d]	−1.09 (0.38)	.02	−2.07 to −0.12
TGT[e]	WLC	−0.35 (0.36)	.81	−1.27 to 0.57
SOL[f]				
Week 4				
ICBT	WLC	−11.73 (3.40)	.003	−20.36 to −3.10
TGT	WLC	−5.88 (3.30)	.27	−14.24 to 2.48
Week 8				
ICBT	WLC	−7.74(3.45)	.10	−16.48 to 1.01
TGT	WLC	−5.00 (3.29)	.42	−13.34 to 3.34
TST[g]				
Week 4				
ICBT	WLC	33.07 (7.07)	<.001	15.17 to 50.98
TGT	WLC	20.70 (6.79)	.01	3.51 to 37.89
Week 8				
ICBT	WLC	17.50 (7.15)	.06	−0.63 to 35.63
TGT	WLC	0.02 (6.77)	.99	−17.13 to 17.17
Sleep efficiency				
Week 4				
ICBT	WLC	2.94 (1.83)	.36	−1.70 to 7.59
TGT	WLC	3.47 (1.76)	.18	−1.00 to 7.94
Week 8				
ICBT	WLC	4.53 (1.85)	.06	−0.18 to 9.25
TGT	WLC	1.57 (1.77)	.85	−2.92 to 6.07
AIS[h]				
Week 4				
ICBT	WLC	−1.62 (0.48)	.004	−2.85 to −0.39
TGT	WLC	−2.27 (0.47)	<.001	−3.46 to −1.08
Week 8				
ICBT	WLC	−0.62 (0.49)	.61	−1.87 to 0.63
TGT	WLC	−0.41 (0.47)	.85	−1.60 to 0.78
GAD-7[i]				
Week 4				
ICBT	WLC	−0.22 (0.40)	.98	−1.24 to 0.80
TGT	WLC	−0.82 (0.39)	.14	−1.81 to 0.16
Week 8				
ICBT	WLC	−0.01 (0.41)	.99	−1.05 to 1.03
TGT	WLC	0.04 (0.39)	.99	−0.95 to 1.03
PHQ-9[j]				

Secondary outcomes and assigned group	Control group	Estimation (SE)	Adjusted *P* value[a]	95% CI (adjusted)
Week 4				
ICBT	WLC	−1.70 (0.47)	.001	−2.90 to −0.51
TGT	WLC	−1.78 (0.45)	<.001	−2.94 to −0.63
Week 8				
ICBT	WLC	−0.65 (0.48)	.53	−1.86 to 0.56
TGT	WLC	−0.41 (0.46)	.84	−1.57 to 0.74
CES-D[k]				
Week 4				
ICBT	WLC	−0.13 (0.47)	.99	−1.32 to 1.06
TGT	WLC	0.02 (0.45)	.99	−1.13 to 1.17
Week 8				
ICBT	WLC	0.98 (0.48)	.16	−0.23 to 2.19
TGT	WLC	1.03 (0.45)	.10	−0.12 to 2.18

[a]Dunnett–Hsu.

[b]PSQI: Pittsburgh Sleep Quality Index.

[c]ICBT: internet-based cognitive behavioral therapy.

[d]WLC: waiting list control.

[e]TGT: Three Good Things.

[f]SOL: sleep onset latency.

[g]TST: total sleep time.

[h]AIS: Athens Insomnia Scale.

[i]GAD-7: Generalized Anxiety Disorder-7.

[j]PHQ-9: Patient Health Questionnaire-9.

[k]CES-D: Center for Epidemiologic Studies–Depression Scale.

At 4 weeks, the adjusted mean change from baseline in the ICBT group compared with the WLC group was −11.73 for SOL (95% CI −20.36 to −3.10; *P*=.003), 17.50 for TST (95% CI 15.17-50.98; *P*<.001), −1.62 for AIS (95% CI −2.85 to −0.39; *P*=.004), and −1.70 for PHQ-9 (95% CI −2.90 to −0.51; *P*=.002). There was a significant improvement in the ICBT group compared with the WLC group in SOL, TST, AIS, and PHQ-9 scores at 4 weeks.

At 4 weeks, the adjusted mean change from baseline in the TGT group compared with the WLC group was 20.70 for TST (95% CI 3.51-37.89; *P*=.01), −2.27 for AIS (95% CI −3.46 to −1.08; *P*<.001), and −1.78 for PHQ-9 (95% CI −2.94 to −0.63; *P*=.001). A significant improvement in the TGT group compared with the WLC group in TST, AIS, and PHQ-9 scores was observed at 4 weeks.

At 8 weeks, the adjusted mean change from baseline in the ICBT group compared with the WLC group was −1.09 for PSQI (95% CI −2.07 to −0.12; *P*=.02). There was a significant improvement in the ICBT group compared with the WLC group in PSQI at 8 weeks.

At 8 weeks, the adjusted mean change from baseline in the TGT group compared with the WLC group did not show a significant difference.

Effectiveness of the Intervention

Table 6 shows the standardized change from baseline to the postintervention survey (at 4 and 8 weeks). Immediately after the intervention, at 4 weeks, the primary outcome effect size (Hedge g) for the PSQI was 0.81 in the ICBT group (95% CI 6.90-8.20) and 0.76 in the TGT group (95% CI 7.06-8.20).

Table 6. Standardized changes of baseline to postintervention outcomes.

Outcomes	Values, mean (SD)			Effect size (95% CI)[a,b]		
	ICBT[c]	TGT[d]	WLC[e]	ICBT	TGT	WLC
PSQI[f]						
Baseline	9.63 (2.32)	9.41 (2.05)	9.45 (2.20)	9.10-10.15	8.97-9.84	9.02-9.88
Week 4	7.55 (2.79)	7.63 (2.62)	8.90 (2.89)	0.81 (6.90-8.20)	0.76 (7.06-8.20)	0.22 (8.30-9.50)
Week 8	7.72 (3.13)	8.17 (2.66)	8.57 (2.93)	0.70 (6.96-8.48)	0.52 (7.59-8.75)	0.34 (7.97-9.18)
SOL[g]						
Baseline	41.75 (27.88)	47.78 (40.64)	41.23 (34.96)	35.50-47.99	39.17-56.39	34.40-48.07
Week 4	23.41 (22.55)	32.61 (27.20)	34.87 (37.29)	0.72 (18.18-28.63)	0.44 (26.70-38.51)	0.18 (27.10-42.63)
Week 8	27.90 (26.44)	33.72 (29.90)	34.43 (32.74)	0.51 (21.50-34.30)	0.39 (27.19-40.25)	0.20 (27.72-41.13)
TST[h]						
Baseline	318.56 (63.66)	323.81 (52.11)	319.93 (61.73)	304.30-332.82	312.77-334.85	307.87-332.00
Week 4	360.84 (60.28)	350.18 (61.20)	323.57 (65.53)	0.68 (346.87-374.80)	0.46 (336.90-363.46)	0.06 (309.92-337.22)
Week 8	355.90 (62.76)	338.67 (58.26)	333.27 (64.46)	0.59 (340.71-371.09)	0.27 (325.95-351.40)	0.21 (320.06-346.47)
Sleep efficiency						
Baseline	82.12 (15.24)	85.36 (10.93)	84.25 (14.68)	78.64-85.61	83.02-87.70	81.34-87.17
Week 4	89.07 (11.77)	90.68 (8.96)	86.76 (13.41)	0.51 (86.26-91.87)	0.53 (88.73-92.64)	0.18 (83.92-89.60)
Week 8	89.79 (9.04)	87.62 (10.86)	86.67 (14.75)	0.60 (87.52-92.07)	0.21 (85.15-90.08)	0.16 (83.61-89.72)
AIS[i]						
Baseline	10.19 (2.95)	9.95 (2.66)	10.24 (2.90)	9.53-10.85	9.39-10.52	9.68-10.81
Week 4	8.14 (3.18)	7.29 (3.67)	9.72 (3.67)	0.67 (7.40-8.87)	0.84 (6.49-8.08)	0.16 (8.97-10.47)
Week 8	8.47 (3.89)	8.47 (3.59)	9.09 (3.28)	0.50 (7.53-9.41)	0.47 (7.69-9.25)	0.37 (8.43-9.76)
GAD-7[j]						
Baseline	2.67 (2.37)	2.91 (2.41)	2.66 (2.37)	2.14-3.20	2.40-3.42	2.20-3.12
Week 4	2.92 (3.29)	2.45 (2.49)	3.02 (3.18)	0.09 (2.16-3.68)	0.19 (1.91-2.99)	0.13 (2.37-3.67)
Week 8	2.84 (3.20)	3.00 (3.03)	2.80 (2.93)	0.06 (2.06-3.61)	0.03 (2.34-3.66)	0.05 (2.20-3.40)
PHQ-9[k]						
Baseline	3.99 (2.23)	4.18 (2.38)	4.08 (2.26)	3.49-4.49	3.68-4.69	3.64-4.52
Week 4	3.68 (2.88)	3.83 (2.94)	5.33 (4.46)	0.12 (3.01 to 4.34)	0.13 (3.20-4.47)	0.36 (4.42-6.24)
Week 8	3.69 (3.29)	4.28 (3.18)	4.29 (3.34)	0.11 (2.90 to 4.49)	0.03 (3.58-4.97)	0.08 (3.59-5.00)
CES-D[l]						
Baseline	7.01 (3.46)	7.10 (3.38)	7.02 (3.74)	6.24-7.79	6.39-7.82	6.29-7.75
Week 4	7.32 (3.81)	7.60 (3.64)	7.55 (3.65)	0.09 (6.44-8.21)	0.14 (6.81-8.38)	0.14 (6.80-8.30)
Week 8	7.41 (3.33)	7.54 (3.67)	6.51 (4.09)	0.12 (6.60-8.22)	0.12 (6.74-8.34)	0.13 (5.67-7.34)

[a]Hedges *g* (effect size; 95% CI lower limit to upper limit of each outcome).

[b]Effect size values for baseline were not applicable; hence, only 95% CI values have been reported.

[c]ICBT: internet-based cognitive behavioral therapy.

[d]TGT: Three Good Things.

[e]WLC: waiting list control.

[f]PSQI: Pittsburgh Sleep Quality Index.

[g]SOL: sleep onset latency.

[h]TST: total sleep time.

[i]AIS: Athens Insomnia Scale.

[j]GAD-7: Generalized Anxiety Disorder-7.

[k]PHQ-9: Patient Health Questionnaire-9.

[l]CES-D: Center for Epidemiologic Studies–Depression Scale.

The secondary outcome effect sizes (Hedge g) were as follows: 0.72 for SOL in the ICBT group (95% CI 18.18-28.63), 0.68 for TST in the ICBT group (95% CI of 346.87-374.80), 0.51 for SE in the ICBT group (95% CI 86.26-91.87) and 0.53 in the TGT group (95% CI 88.73-92.64), and 0.67 for AIS in the ICBT group (95% CI 7.40-8.87) and 0.84 in the TGT group (95% CI 6.49-8.08).

After the follow-up period, at 8 weeks, the secondary outcome effect sizes (Hedges *g*) were as follows: 0.70 for PSQI in the ICBT group (95% CI 6.96-8.48) and 0.52 in the TGT group (95% CI 7.59-8.75), 0.51 for SOL in the ICBT group (95% CI 21.50-34.30), 0.59 for TST in the ICBT group (95% CI 340.71-371.09), 0.60 for SE in the ICBT group (95% CI 87.52-92.07), and 0.50 for AIS in the ICBT group (95% CI 7.53-9.41).

Discussion

Principal Findings

The results of this study suggest that a self-help internet intervention without email support is effective as a noninvasive self-medication for adults with insomnia. The ICBT group showed a significant change of −6.11 in the PSQI score. The participant characteristics of this study differed significantly from those of similar previous studies. A previous study by van Straten et al [22] showed the efficacy of guided ICBT using an RCT design, in which 1500 patients with insomnia were invited by email, and 118 patients were assigned to the ICBT group (n=59) and a WLC group (n=59). However, the reported baseline mean PSQI score in the ICBT group in that study was relatively severe (12.4, SD 2.1). The mean change in PSQI scores in the ICBT group was −3.5, and in the WLC group it was −0.1. Our previous study of ICBT guided by email support reported a severe baseline PSQI mean score of 13.5 (SD 2.7); furthermore, participants had persistent symptoms despite taking sleep medications such as benzodiazepines [9].

In this study, the unguided ICBT group showed a change of −2.19, the TGT group showed a change of −1.79, and the WLC group showed a change of −0.64; the mean change in PSQI scores was smaller than those reported by van Straten et al [22] and Sato et al [9]. This could be because of the baseline average score of PSQI being 9.8, which indicates relatively mild insomnia. In addition, given that the ICBT was unguided, participant motivation to continue in this study may not have been well-supported.

The TGT group showed a significant decrease in PSQI scores compared with the WLC group only at the end of the program, that is, at the 4-week time point. To the best of our knowledge, no previous studies have examined the effects of TGT on insomnia through an RCT. However, as no significant improvement in positive emotions (indicated by the actual items on the CES-D) was observed in the TGT group, this casts doubt on the existence of a mechanism for the improvement of insomnia through the improvement of positive emotions. Contrary to previous studies, the TGT exercise did not lead to the improvement of positive emotions, and thus, it is unclear whether TGT really contributes to the improvement of happiness.

In this study, comparing the ICBT group and the TGT group was not the original purpose; however, the number of participants who received >75% of the intervention in the ICBT group was 70% (55/79) of participants compared with 90% (79/88) of participants who received >75% of the intervention in the TGT group. This could have been because of difficulties with regular participation in the ICBT program.

At the 4- and 8-week follow-ups, the ICBT group showed a significant improvement in PSQI scores over the WLC group, whereas the TGT group did not show a significant improvement over the WLC group. Therefore, ICBT could have longer-term effectiveness. If further research shows that ICBT results in sustained effects on insomnia, and TGT is easier to participate in more regularly, it may be necessary to consider an unguided intervention that combines both ICBT and TGT to achieve optimal outcomes.

Limitations

In this study, there were errors in the calculation set by the internet research company that conducted the web questionnaire for PSQI, PHQ-9, and GAD-7 scores, which may have affected the eligible participants. In future studies, the distribution of inclusion and exclusion criteria should be carefully monitored to avoid similar mistakes.

Some participants in the ICBT and TGT groups did not participate more than once. However, in the WLC group, all participants were included in the final analysis. To ensure consistency with the ICBT and TGT groups in future research, it will be necessary to select those who have participated at least once in the WLC group for inclusion in the final analysis, or alternatively a placebo sham program could be implemented.

Conclusions

In conclusion, this RCT provided evidence that 4 weeks of unguided ICBT and TGT exercise for adults with insomnia, both administered as self-help internet interventions without email support, may improve insomnia symptoms compared with the WLC group. The findings also suggest that the effects of ICBT may last longer than TGT, whereas TGT may be easier to participate in more regularly. However, this study experienced errors during the selection process. Further research is warranted to examine the effectiveness of a combined intervention of ICBT and TGT for insomnia.

Acknowledgments

The authors thank all the participants for their valuable contributions to this study. This study was funded by the Research Institute of Economy, Trade, and Industry (RIETI), Japan, and supported, in part, by the Grant-in-Aid for Scientific Research (B) from the Japan Society for the Promotion of Science (grant KAKENHI 17H04091 to ES), the Grant-in-Aid for Scientific Research (C) from Japan Society for the Promotion of Science (grant KAKENHI 18K10716 to DS), and grants from the Chiba Foundation for Health Promotion and Disease Prevention. The RIETI was involved in the conduct of the study through its contact with the market research company. The RIETI had no role in the study design, data analysis, decision to publish, or preparation of the manuscript.

Authors' Contributions

DS, YS, CS, YH, SO, MH, and ES contributed to the design of the study. RT performed statistical analysis. DS, YS, and ES interpreted the data and drafted the manuscript. All authors have approved the manuscript.

Conflicts of Interest

SD and ES have developed the web-based treatment but have no commercial interests.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 935 KB - jmir_v24i2e28747_app1.pdf]

References

1. Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). Washington, D.C: American Psychiatric Association; 2013.
2. Ohayon MM. Epidemiology of insomnia: what we know and what we still need to learn. Sleep Med Rev 2002 Apr;6(2):97-111. [doi: 10.1053/smrv.2002.0186] [Medline: 12531146]
3. Komada Y, Nomura T, Kusumi M, Nakashima K, Okajima I, Sasai T, et al. Correlations among insomnia symptoms, sleep medication use and depressive symptoms. Psychiatry Clin Neurosci 2011 Feb;65(1):20-29 [FREE Full text] [doi: 10.1111/j.1440-1819.2010.02154.x] [Medline: 21105960]
4. Kim K, Uchiyama M, Okawa M, Liu X, Ogihara R. An epidemiological study of insomnia among the Japanese general population. Sleep 2000 Feb 01;23(1):41-47. [Medline: 10678464]
5. Hafner M, Stepanek M, Taylor J, Troxel WM, van Stolk C. Why sleep matters-the economic costs of insufficient sleep: a cross-country comparative analysis. Rand Health Q 2017 Jan;6(4):11 [FREE Full text] [Medline: 28983434]
6. Yamauchi M. [Diagnosis of and therapy for sleep disorders in primary care]. Seishin Shinkeigaku Zasshi 2006;108(11):1230-1236. [Medline: 17375456]
7. Seligman ME, Steen TA, Park N, Peterson C. Positive psychology progress: empirical validation of interventions. Am Psychol 2005;60(5):410-421. [doi: 10.1037/0003-066X.60.5.410] [Medline: 16045394]
8. Sato D, Yoshinaga N, Nagai E, Hanaoka H, Sato Y, Shimizu E. Randomised controlled trial on the effect of internet-delivered computerised cognitive-behavioural therapy on patients with insomnia who remain symptomatic following hypnotics: a study protocol. BMJ Open 2018 Jan 30;8(1):e018220 [FREE Full text] [doi: 10.1136/bmjopen-2017-018220] [Medline: 29382675]
9. Sato D, Yoshinaga N, Nagai E, Nagai K, Shimizu E. Effectiveness of internet-delivered computerized cognitive behavioral therapy for patients with insomnia who remain symptomatic following pharmacotherapy: randomized controlled exploratory trial. J Med Internet Res 2019 Apr 11;21(4):e12686 [FREE Full text] [doi: 10.2196/12686] [Medline: 30973344]
10. Ong AD, Kim S, Young S, Steptoe A. Positive affect and sleep: a systematic review. Sleep Med Rev 2017 Oct;35:21-32. [doi: 10.1016/j.smrv.2016.07.006] [Medline: 27553441]
11. Seligman M. Flourish: a new understanding of happiness and wellbeing: the practical guide to using positive psychology to make you happier and healthier. Hachette, UK: Nicholas Brealey Publishing; 2011.
12. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
13. Doi Y, Minowa M, Uchiyama M, Okawa M, Kim K, Shibui K, et al. Psychometric assessment of subjective sleep quality using the Japanese version of the Pittsburgh Sleep Quality Index (PSQI-J) in psychiatric disordered and control subjects. Psychiatry Res 2000 Dec 27;97(2-3):165-172. [doi: 10.1016/s0165-1781(00)00232-8] [Medline: 11166088]
14. Soldatos CR, Dikeos DG, Paparrigopoulos TJ. Athens Insomnia Scale: validation of an instrument based on ICD-10 criteria. J Psychosom Res 2000 Jun;48(6):555-560. [doi: 10.1016/s0022-3999(00)00095-7] [Medline: 11033374]
15. Okajima I, Nakajima S, Kobayashi M, Inoue Y. Development and validation of the Japanese version of the Athens Insomnia Scale. Psychiatry Clin Neurosci 2013 Sep;67(6):420-425 [FREE Full text] [doi: 10.1111/pcn.12073] [Medline: 23910517]

16. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
17. Doi S, Ito M, Takebayashi Y, Muramatsu K, Horikoshi M. Factorial Validity and Invariance of the 7-Item Generalized Anxiety Disorder Scale (GAD-7) Among Populations With and Without Self-Reported Psychiatric Diagnostic Status. Front Psychol 2018 Sep 19;9:1741-1749 [FREE Full text] [doi: 10.3389/fpsyg.2018.01741] [Medline: 30283386]
18. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
19. Muramatsu K, Miyaoka H, Kamijima K, Muramatsu Y, Yoshida M, Otsubo T, et al. The patient health questionnaire, Japanese version: validity according to the mini-international neuropsychiatric interview-plus. Psychol Rep 2007 Dec;101(3 Pt 1):952-960. [doi: 10.2466/pr0.101.3.952-960] [Medline: 18232454]
20. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. Applied Psychol Measurement 1977;1(3):385-401.
21. Shima S, Shikano T, Kitamura T, Asai M. New self-rating scale for depression. Clin Psychiatry 1985;27(6):717-723 [FREE Full text]
22. van Straten A, Emmelkamp J, de Wit J, Lancee J, Andersson G, van Someren EJ, et al. Guided internet-delivered cognitive behavioural treatment for insomnia: a randomized trial. Psychol Med 2014 May;44(7):1521-1532. [doi: 10.1017/S0033291713002249] [Medline: 24001364]

Abbreviations

AIS: Athens Insomnia Scale
CBT: cognitive behavioral therapy
CES-D: Center for Epidemiologic Studies–Depression Scale
CONSORT: Consolidated Standards of Reporting Trials
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
FAS: full analysis set
GAD-7: Generalized Anxiety Disorder-7
ICBT: internet-based cognitive behavioral therapy
PHQ-9: Patient Health Questionnaire-9
PPS: per-protocol set
PSQI: Pittsburgh Sleep Quality Index
RCT: randomized controlled trial
RIETI: Research Institute of Economy, Trade, and Industry
SE: sleep efficiency
SOL: sleep onset latency
TGT: Three Good Things
TST: total sleep time
UC: usual care
WLC: waiting list control

Edited by A Mavragani; submitted 17.03.21; peer-reviewed by Y Xu, E Kawamoto; comments to author 27.04.21; revised version received 11.05.21; accepted 12.12.21; published 09.02.22.

Please cite as:
Sato D, Sekizawa Y, Sutoh C, Hirano Y, Okawa S, Hirose M, Takemura R, Shimizu E
Effectiveness of Unguided Internet-Based Cognitive Behavioral Therapy and the Three Good Things Exercise for Insomnia: 3-Arm Randomized Controlled Trial
J Med Internet Res 2022;24(2):e28747
URL: https://www.jmir.org/2022/2/e28747
doi:10.2196/28747
PMID:35138259

Original Paper

Web-Based Cognitive Testing in Psychiatric Research: Validation and Usability Study

Amy Joanne Lynham[1], BSc, PhD; Ian R Jones[1], BSc, MBBS, MSc, PhD; James T R Walters[1], BM, MSc, PhD

Medical Research Council Centre for Neuropsychiatric Genetics and Genomics, Division of Psychiatry and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, United Kingdom

Corresponding Author:
James T R Walters, BM, MSc, PhD
Medical Research Council Centre for Neuropsychiatric Genetics and Genomics
Division of Psychiatry and Clinical Neurosciences, School of Medicine
Cardiff University
Cardiff, CF24 4HQ
United Kingdom
Phone: 44 29206 88434
Email: waltersjt@cardiff.ac.uk

Abstract

Background: Cognitive impairments are features of many psychiatric disorders and affect functioning. A barrier to cognitive research on psychiatric disorders is the lack of large cross-disorder data sets. However, the collection of cognitive data can be logistically challenging and expensive. Web-based collection may be an alternative; however, little is known about who does and does not complete web-based cognitive assessments for psychiatric research.

Objective: The aims of this study are to develop a web-based cognitive battery for use in psychiatric research, validate the battery against the Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) Consensus Cognitive Battery, and compare the characteristics of the participants who chose to take part with those of the individuals who did not participate.

Methods: Tasks were developed by The Many Brains Project and selected to measure the domains specified by the MATRICS initiative. We undertook a cross-validation study of 65 participants with schizophrenia, bipolar disorder, depression, or no history of psychiatric disorders to compare the web-based tasks with the MATRICS Consensus Cognitive Battery. Following validation, we invited participants from 2 large ongoing genetic studies, which recruited participants with psychiatric disorders to complete the battery and evaluated the demographic and clinical characteristics of those who took part.

Results: Correlations between web-based and MATRICS tasks ranged between 0.26 and 0.73. Of the 961 participants, 887 (92.3%) completed at least one web-based task, and 644 (67%) completed all tasks, indicating adequate completion rates. Predictors of web-based participation included being female (odds ratio [OR] 1.3, 95% CI 1.07-1.58), ethnicity other than White European (OR 0.66, 95% CI 0.46-0.96), higher levels of education (OR 1.19, 95% CI 1.11-1.29), diagnosis of an eating disorder (OR 2.17, 95% CI 1.17-4) or depression and anxiety (OR 5.12, 95% CI 3.38-7.83), and absence of a diagnosis of schizophrenia (OR 0.59, 95% CI 0.35-0.94). Lower performance on the battery was associated with poorer functioning (B=−1.76, SE 0.26; P<.001).

Conclusions: Our findings offer valuable insights into the advantages and disadvantages of testing cognitive function remotely for mental health research.

(J Med Internet Res 2022;24(2):e28233) doi:10.2196/28233

KEYWORDS

cognition; mental health; online; digital; assessment; validation; memory; attention; mobile phone

Introduction

Background

Cognitive impairments are core features of many psychiatric disorders [1-3], persist during remission [2,3], are not alleviated by current treatments [4-6], and are associated with poor functional outcomes [7,8]. Existing cognitive research on psychiatric disorders is limited by relatively small sample sizes, as collecting cognitive data can be labor intensive, and the use of existing cognitive batteries can be expensive given licensing restrictions. A potential solution to this is to use web-based data

collection methods, which may be an effective way of acquiring large amounts of cognitive data using minimal resources. Data can be collected remotely using specially designed tasks that are administered via the internet and accessed using the participants' own devices and web browsers in an unsupervised setting, such as the participants' homes. The advantages of web-based methods include (1) relatively inexpensive costs per participant [9,10], (2) automatic data entry that limits errors [10], (3) ability to recruit from locations that would normally be out of reach [11], and (4) promotion of research to the public [11].

To date, studies comparing web-based and laboratory-based cognitive tasks have reported high correlations, few systematic differences between the assessments, and good internal reliability of web-based tasks [9,12-15]. One such study is the global citizen science project, *TestMyBrain*, which has collected cognitive data via the internet on >2 million volunteers around the world. Analysis of the data collected from the website compared with data collected using the same tasks under supervision in a research laboratory showed few differences in mean performance, variance, and internal reliability between the 2 sets of tasks [12]. A recent study of UK Biobank (UKB) participants demonstrated adequate concurrent validity between 11 cognitive assessments administered without supervision and 11 previously validated tests from the published literature (correlations between 0.22 and 0.83) [16]. The highest correlation (r=0.83) was observed between versions of the same task—the National Institutes of Health Toolbox Picture Vocabulary Test—whereas the lowest correlations were observed between tasks with differing methodologies (UKB Pairs Matching test and Wechsler Memory Scale-IV Designs Total; r=−0.33) or tasks with low levels of performance variance (UKB Prospective Memory test and Rivermead Behavioral Memory Test Appointments; r=0.22). Similar results were reported in a study comparing the Amsterdam Cognition Scan with a traditional test battery (correlations ranged between 0.36 and 0.78), where lower correlations were observed in tasks that had differing designs [15]. Another study compared identical Cambridge Neuropsychological Test Automated Battery tasks administered unsupervised via the internet and administered in person at a research facility (r=0.39-0.73) [14]. Correlations were similar to previously reported test–retest reliabilities for the tasks, although it was noted that tasks with a reaction time component were less comparable across the different administration settings.

To date, studies have evaluated the use of web-based cognitive assessments in healthy population samples. Therefore, questions remain about whether web-based tasks conducted in an unsupervised setting without a researcher present are suitable for research on individuals with psychiatric disorders. This is a particularly important question, given that these individuals are more likely to have moderate to severe cognitive impairments. An issue is participation bias, as web-based studies may exclude individuals who are less computer literate or who do not have internet access, such as older adults, those with lower incomes or education levels, or those with more severe psychiatric disorders. For example, a study that used Facebook to recruit participants for a mental health survey found that the participants were younger, more likely to be female, more educated, and more likely to be English speakers compared with the national averages taken from census data and a population study [17].

Objectives

This study has 2 aims. Our first aim is to develop and validate a web-based cognitive battery for use in psychiatric research. Our second aim is to determine whether web-based cognitive testing is a suitable method for large-scale mental health research. By offering participation in a web-based cognitive study to those who had already been recruited into a cohort of individuals with mental health conditions, we will be able to identify the characteristics of those who chose to take part and compare them with those who did not. This study is presented in two parts: (1) validation of the web-based battery and (2) expansion of the web-based battery via recruitment of a large cross-disorder sample.

Methods

Participants

Participants from 2 studies conducted within the Medical Research Council Centre for Neuropsychiatric Genetics and Genomics at Cardiff University were invited to take part: (1) the National Centre for Mental Health (NCMH) cohort study [18] and (2) the Cognition in Mood, Psychosis, and Schizophrenia Study (CoMPaSS) [19]. Diagnoses were ascertained through either self-report in the NCMH or through a clinical interview in the CoMPaSS (Schedule for Clinical Assessment in Neuropsychiatry [20]). In the NCMH, participants were asked the following question: "Has a doctor or health professional ever told you that you have any of the following diagnoses?" Participants were given a list of diagnoses and asked to indicate all diagnoses that applied, the diagnosis they considered to be their primary diagnosis, and whether their clinical team would agree. Both studies included confirmation of consent from participants to be approached for other research within the Centre and consent for medical records to be accessed to obtain information regarding diagnosis and other clinical details. The full details of these studies can be found in Multimedia Appendix 1 [12,19-39]. All stages of the web-based study received ethical approval from the School of Medicine research ethics committee at Cardiff University (reference number 15/64), and each parent study had National Health Service ethical approval (NCMH reference number: 16/WA/0323; CoMPaSS reference number: 07/WSE03/110). The participants indicated their consent on the web by ticking a box at the bottom of the information page. We did not collect any identifiable information about the participants during the assessments; the participants' data were pseudonymized and linked with an ID number. The assessment websites complied with current UK data security best practice guidelines in consultation with the Information Technology Systems Security Team and Research Governance Officers at Cardiff University.

Neuropsychological Assessments

Web-Based Battery

All tasks were selected from and hosted on The Many Brains Project's web-based cognitive testing platform, TestMyBrain [12,40]. We selected the tasks to assess, as closely as possible, the domains outlined by the National Institute of Mental Health–Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) initiative [41]. We were able to include a formally equivalent web-based version of digit symbol coding. However, it was not possible to select formally equivalent tasks for the remaining domains because of a lack of availability of web-based versions. In addition to these domains, we included a measure of crystallized intelligence and a measure of risk-taking propensity. The final battery included 9 tasks (see Table 1 for domains, tasks, and equivalent MATRICS Consensus Cognitive Battery [MCCB] tasks). Full descriptions of the tasks can be found in Multimedia Appendix 1. The total administration time was approximately 45 to 50 minutes.

The tests were designed to run on desktop and laptop computers, touchscreen tablet computers, and smartphones. The Many Brains Project developed the cognitive tasks and hosted them on their secure TestMyBrain server, which could be accessed using a study-specific website link. The assessments were loaded in the participant's internet browser, and the data were stored locally during each task. At the end of each task, the data were encrypted and uploaded to a secure server.

Table 1. Web-based and Measurement and Treatment Research to Improve Cognition in Schizophrenia Consensus Cognitive Battery (MCCB) tasks[a].

Domain	MCCB task	Web-based task
Speed of processing	BACS[b]: digit symbol coding	Digit symbol coding
Social cognition	MSCEIT[c]: Managing Emotions	Morphed emotion identification
Verbal learning	Hopkins Verbal Learning Test–Revised	Verbal paired associates
Working memory	Letter–number sequencing	Backward digit span
Visual learning	Brief Visuospatial Memory Test–Revised	Hartshorne visual working memory
Reasoning and problem solving	NAB[d]: Mazes	Matrix Reasoning Test
Strategic risk taking[e]	N/A[f]	Balloon Analogue Risk Task
Attention	Continuous Performance Test–Identical Pairs	Multiple object tracking
Premorbid IQ[g]	National Adult Reading Test–Revised	Vocabulary

[a]Web-based tasks are shown in order of administration.

[b]BACS: Brief Assessment of Cognition in Schizophrenia.

[c]MSCEIT: Mayer–Salovey–Caruso Emotional Intelligence Test.

[d]NAB: Neuropsychological Assessment Battery.

[e]No equivalent offline measure was included.

[f]N/A: not applicable.

[g]No equivalent MCCB task; thus, the National Adult Reading Test was included for comparison.

Reference Battery

The MCCB was administered to participants as the reference battery in part 1 of the study to validate the web-based battery. The MCCB was created through the MATRICS initiative with the explicit aim of developing a consensus cognitive battery that could be used in schizophrenia research. The selection of tasks for the MCCB was driven by expert panels; consultations with scientists; evaluations of factor-analytic studies to identify relevant domains; and assessments of the psychometric properties, practicality, and tolerability of existing cognitive tasks [41,42]. The final MCCB comprises 10 tasks assessing 7 domains of cognition. In addition to the MCCB, the National Adult Reading Test (NART) was administered as a measure of premorbid IQ [21]. The NART is a measure of vocabulary that comprises 50 irregularly spelled words that the participant must read aloud, and it was included as a reference test for the web-based vocabulary test.

Clinical and Demographic Variables

In addition to the cognitive assessment, participants answered questions about current diagnosis, medication, education, occupation, and current mood. The 12-item self-report version of the World Health Organization Disability Assessment Schedule Version 2 (WHODAS 2.0) was included as a measure of functioning [43]. The web-based questionnaire also included the Hospital Anxiety and Depression Scale [44] and the Altman Self-Rating Mania Scale [45]. Data on lifetime diagnosis, age of onset, and hospital admissions were obtained from the parent studies, CoMPaSS and NCMH (Multimedia Appendix 1).

Part 1: Validation Study

Participants

Participants with major depressive disorder (15/65, 23%), bipolar disorder (type 1: 11/65, 17%; type 2: 5/65, 8%), or schizophrenia (15/65, 23%), as well as healthy controls (19/65, 29%) were recruited for the validation study. We selected these 3 diagnostic groups based on extensive research establishing

the characteristics of cognitive performance of participants with each of these disorders using offline, traditional cognitive testing. For this study, we decided upon a conservative definition of depression that required a reported diagnosis and previous treatment with at least one antidepressant medication. This definition has been used in a study using self-reported measures of depression in the UKB [46]. Informed consent was obtained at both stages, in writing before administration of the MCCB and on the web before completing the web-based cognitive battery, and the participants were reimbursed for their participation.

Study Design

The participants were asked to complete two cognitive batteries on consecutive days: (1) the MCCB and the NART [21] and (2) the web-based battery (Table 1). A trained researcher administered the MCCB in a supervised setting on the first day, and then the participants were asked to complete the web-based battery unsupervised on the second day. The order of completion was not counterbalanced for practical reasons. First, some participants from the NCMH were recruited prospectively by completing the MCCB with a researcher immediately after completing the NCMH assessment; therefore, we could not randomly assign the order of completion. Second, it would have been difficult to ensure that the participants assigned to complete the web-based part first did so before their appointment as they were not supervised. The participants were asked to return a feedback questionnaire on completion of the study. They rated the overall web-based battery on enjoyment, duration, and difficulty and rated the clarity of the given instructions and information. They named the tasks that they liked the most and the least. They were also asked to provide information on any technical difficulties they experienced.

Data Analysis

All statistical analyses were conducted using R (version 3.3.0; R Foundation for Statistical Computing). Convergent validity was examined by conducting correlations between the MCCB and web-based tasks that assessed equivalent cognitive domains (Table 1). Correlations were conducted across the entire sample and for cases only. Pearson correlations were used when task performance was normally distributed, whereas Spearman correlations were used for tasks that were not normally distributed. A correlation matrix of all web-based tasks and the MCCB-equivalent tasks was also generated. Partial correlations were used to adjust for the time between completion of the batteries, age, and *g* (excluding domain of interest). Finally, correlation analyses were repeated after stratification by input device type (keyboard or touchscreen). Although previous studies do not appear to have corrected for multiple testing as the correlation coefficients are more important for validation [14-16], we calculated the corrected *P* values using the false discovery rate method.

Part 2: Feasibility of Web-Based Cognitive Testing

Participants

Following completion of the validation study, participants from the NCMH and CoMPaSS were sent invitation letters or emails with instructions on how to participate in the study and their unique website link. A reminder was sent to participants who did not respond to our initial invitation. Response rates can be found in Multimedia Appendix 1. The collected data were combined with the validation data set. Participants were excluded from the analyses if they reported a neurological condition likely to affect cognitive function or if they did not complete any of the cognitive tasks. Diagnostic groups with >20 participants were included in the analyses. Healthy controls were excluded from the analyses if they reported a history of psychiatric diagnosis or medication or first-degree family history of schizophrenia, bipolar disorder, autism, or intellectual disability. Therefore, the final sample for analysis (N=887) included 21.1% (187/887) of controls, 16.5% (146/887) of participants with bipolar spectrum disorders, 4.8% (43/887) of participants with schizophrenia, 29.4% (261/887) of participants with unipolar depression, 7.6% (67/887) of participants with anxiety disorders, 5.4% (48/887) of participants with posttraumatic stress disorder, 2.4% (21/887) of participants with an eating disorder, and 12.9% (114/887) of participants who reported comorbid depression and anxiety disorders. The flow diagram in Multimedia Appendix 1 shows a breakdown of the number of participants excluded according to our criteria.

Data Analysis

Preparation of Cognitive Data

For each task, z scores were derived using the mean and SD of the controls (187/887, 21.1%). General cognitive performance *g* was derived using multidimensional scaling (MDS) for participants who had completed at least five tasks, as we have done in a previous study [47]. MDS is an analogous approach to principal component analysis; however, an advantage of this approach is that it can accommodate missing data [48,49]. General cognitive performance *g* was calculated as the first dimension produced by the MDS analysis. Measures of *g* derived from MDS and principal component analysis were highly correlated (r=0.996). All statistical analyses were conducted using R (version 3.3.0).

Sample Characteristics

We compared those who participated in part 2 of the web-based study (n=1152) with those who were invited but did not take part (nonresponders; n=5768) to assess whether there was recruitment bias in the web-based sample. These comparisons were conducted separately for cases and controls and were further separated by original study (NCMH or CoMPaSS). Logistic regressions were conducted with participation in the study as the outcome and the following predictors: age; sex; education; lifetime occupation; ethnicity; time since recruitment into parent study; and, among cases only, diagnosis, age of illness onset, and ever admitted to a psychiatric hospital.

Cognitive Performance and Functioning

We performed linear regressions with cognitive score as the predictor, total score on the WHODAS 2.0 as the outcome, and age and sex as covariates to test the association between cognitive performance and functioning. We repeated this regression by covarying for diagnosis. We ran separate linear regressions for each cognitive task and *g*. *P* values were corrected using the false discovery rate method.

Comparing Cognition Between Diagnostic Groups

We compared cognitive performance between healthy controls (187/887, 21.1%), major depressive disorder (295/887, 33.3%), bipolar disorder (116/887, 13.1%), and schizophrenia (38/887, 4.3%) using analysis of covariance with age and sex as covariates. Owing to the self-report nature of diagnoses in the NCMH, more conservative inclusion criteria were used to define depression, bipolar disorder, and schizophrenia in these analyses by considering medication use (see Multimedia Appendix 1 for details). Analyses of covariance were followed up with the honestly significant difference test by Tukey for pairwise comparisons. Hedge *g* effect sizes were calculated by dividing the mean group difference by the pooled SD [50].

Statement of Ethical Approval

All participants in part 1 (validation study) provided written informed consent. All participants in part 2 (web-based only) were required to indicate their informed consent by selecting *yes* in response to the statement "I agree to take part in this study and know that I am free to leave the study at any point," located on the information page of the study website. All stages of the web-based study titled *Cognition in Mood, Psychosis and Schizophrenia Study (CoMPaSS Web)* received ethical approval from the School of Medicine research ethics committee at Cardiff University (reference number: 15/64). The NCMH received a favorable ethical opinion from Wales research ethics committee 2 (reference: 16/WA/0323). CoMPaSS received a favorable ethical opinion from the South-East Wales research ethics committee panel (reference: 07/WSE03/110). All experiment protocols were conducted in accordance with the ethical standards of the Cardiff University School of Medicine research ethics committee and with the Declaration of Helsinki.

Data Availability

The data sets generated and analyzed during this study are available at the Medical Research Council Centre for Neuropsychiatric Genetics and Genomics Walters Group Data Repository [51].

Results

Part 1: Validation Study

Data Completion and Sample Characteristics

Approximately 89% (58/65) of participants completed all the web-based tasks, and the same number completed all the MCCB tasks and NART (see Table 2 for individual domains). The sample had a wide age range (22 to 78, mean 47, SD 14.8 years) and a higher percentage of women (38/65, 58%) than men. More than one-third of the sample had an undergraduate degree (26/65, 40%; see Multimedia Appendix 1).

Table 2. Results of correlation analyses between web-based and offline tasks that assessed equivalent domains (N=65).

Domain	Participants, n (%)	Cases and controls		Cases only		Partial correlations[a], *r*			Device, *r*	
		r (95% CI)	*P* value	*r* (95% CI)	*P* value	Time	Age	General cognitive performance (*g*)	Keyboard	Touchscreen
Speed of processing	65 (100)	0.73 (0.59 to 0.83)	<.001	0.69 (0.50 to 0.82)	<.001	0.74	0.66	0.39	0.75[b]	0.75[b]
Verbal learning[c]	63 (97)	0.41 (0.18 to 0.57)	.002	0.40 (0.11 to 0.64)	.02	0.42	0.36	0.19	0.48[b]	0.24
Working memory	64 (98)	0.34 (0.10 to 0.54)	.01	0.36 (0.07 to 0.59)	.03	0.33	0.31	0.1	0.3	0.43
Visual learning	63 (97)	0.12 (−0.13 to 0.36)	.35	0.15 (−0.15 to 0.43)	.33	0.12	0.02	−0.12	0.17	−0.01
Social cognition	63 (97)	0.26 (0.01 to 0.47)	.045	0.33 (0.04 to 0.56)	.04	0.26	0.24	0.1	0.17	0.48
Reasoning and problem solving[c]	64 (98)	0.53 (0.33 to 0.70)	<.001	0.55 (0.29 to 0.75)	<.001	0.53	0.41	0.26	0.55[b]	0.54
Attention	61 (94)	0.34 (0.09 to 0.55)	.01	0.31 (−0.01 to 0.56)	.06	0.34	0.27	0.08	0.36[b]	0.29
Premorbid IQ[c]	62 (95)	0.64 (0.44 to 0.78)	<.001	0.60 (0.38 to 0.76)	<.001	0.64	0.64	0.59	0.65[b]	0.66[b]
General cognitive performance (*g*)	65 (100)	0.78 (0.66 to 0.86)	<.001	0.79 (0.64 to 0.88)	<.001	0.78	0.72	N/A[d]	0.75[b]	0.85[b]

[a]Correlation coefficients after adjusting for time between completion of the 2 batteries in days, age, and *g* (cases and controls).

[b]Correlations significant after correction for multiple testing.

[c]Spearman rank correlation ρ shown instead because of nonnormal distribution for these tests.

[d]N/A: not applicable.

Convergent Validity

The results of correlations between tasks measuring equivalent domains are shown in Table 2 (performance on the MCCB and web-based battery can be found in Multimedia Appendix 1). The measures of *g* derived from the MCCB and web-based batteries were correlated (r=0.78). Scores from 88% (7/8) of the web-based tasks were correlated with scores from the MCCB equivalents (r=0.26-0.73). These results did not change when the time between completion of the batteries was added as a covariate, although correlation coefficients were attenuated after adjustment for age and *g*. When analyses were restricted to cases with a psychiatric diagnosis, scores from 75% (6/8) of the web-based tasks were correlated with scores from the MCCB equivalents (r=0.33-0.69). Similar correlations were observed when analyses were stratified by input device type (keyboard users: 47/65, 72%, r=0.17-0.75; touchscreen users: 17/65, 26%, r=−0.01 to 0.75); however, fewer correlations were significant because of reduced power. Scores on the Balloon Analogue Risk Task (BART) showed low correlations with the MCCB tasks (r=0.07-0.25; Figure 1).

Figure 1. Pearson correlations between web-based and Measurement and Treatment Research to Improve Cognition in Schizophrenia Consensus Cognitive Battery (MCCB) tasks. Red squares indicate tasks assessing the same domain. Only tasks from the MCCB with an equivalent web-based task are shown (the Trail Making Test-A, Animal Naming Test, and Wechsler Memory Scale-III: Spatial Span were excluded). BVMT-R: Brief Visuospatial Memory Test–Revised; CPT-IP: Continuous Performance Test–Identical Pairs; HVLT-R: Hopkins Verbal Learning Test–Revised; MSCEIT-ME: Mayer–Salovey–Caruso Emotional Intelligence Test–Managing Emotions; NAB: Neuropsychological Assessment Battery; NART: National Adult Reading Test.

Tolerability

Feedback questionnaires were received from 63% (41/65) of participants. Of those who responded, all participants agreed that the instructions given at the start of the study were clear and rated the clinical questionnaire positively. Overall, the cognitive tasks were rated as enjoyable by 100% (41/41) of the participants who responded and as being of reasonable duration and difficulty by 95%. Approximately 3% (2/65) of participants rated the duration and difficulty of the battery as *poor*. The most popular task was multiple object tracking, and the least popular was verbal paired associates. Of the 40 participants who responded to the question, 34 (85%) reported that they would be *more likely* to take part in future web-based studies after taking part in this study, whereas 5 (13%) responded *don't know*, and 1 (3%) responded that they were *less likely*. Of the 40 participants, 6 (15%) reported technical difficulties and, of these 6 participants, 5 (83%) were able to complete all the tasks; thus, this did not affect the availability of data for these participants.

Part 2: Feasibility of Web-Based Cognitive Testing

Sample Characteristics

We compared the demographic and clinical characteristics of all participants recruited during part 2 of the study (n=1152) and nonresponders (n=5768; see Table 3 for cases and Table 4 for controls) to evaluate whether recruitment bias was present in the web-based sample. In CoMPaSS, web-based participants were more highly educated than nonresponders (odds ratio [OR] 1.49, 95% CI 1.15-1.98; P=.004). Among NCMH cases, the

significant predictors for web-based participation were diagnosis of eating disorder (OR 2.17, 95% CI 1.17-4; *P*=.01), diagnosis of comorbid depression and anxiety (OR 5.12, 95% CI 3.38-7.83; *P*<.001), absence of diagnosis of schizophrenia (OR 0.59, 95% CI 0.35-0.94; *P*=.03), older age (OR 1.01, 95% CI 1-1.02; *P*=.003), being female (OR 1.3, 95% CI 1.07-1.58; *P*=.009), ethnicity other than White European (OR 0.66, 95% CI 0.46-0.96; *P*=.03), younger age of onset (OR 0.99, 95% CI 0.98-0.99; *P*=.001), higher level of education (OR 1.19, 95% CI 1.11-1.29; *P*<.001), and shorter time since recruitment into the NCMH (OR 0.98, 95% CI 0.97-0.98; *P*<.001). Among the controls, older age (OR 1.03, 95% CI 1.01-1.04; *P*<.001), higher levels of education (OR 1.5, 95% CI 1.19-1.83; *P*<.001), and shorter time since recruitment into the NCMH (OR 0.95, 95% CI 0.93-0.96; *P*<.001) were associated with web-based participation.

Table 3. Characteristics of participants and nonresponders (cases; N=5981)[a].

Characteristic	NCMH[b] cases				CoMPaSS[c] cases			
	Took part (n=906)	Did not take part (n=4341)	OR[d] (95% CI)	*P* value	Took part (n=33)	Did not take part (n=701)	OR[d] (95% CI)	*P* value
Age (years), mean (SD)	47.82 (14.6)	47.45 (14.77)	1.01 (1-1.02)	.003	50.33 (12.63)	52.19 (12.42)	0.99 (0.96-1.03)	.71
Women, n (%)	668 (73.7)	2838 (65.4)	1.3 (1.07-1.58)	.009	18 (54.5)	283 (40.4)	1.47 (0.68-3.17)	.32
Ethnicity, n (%)			0.66 (0.46-0.96)	.03			0.25 (0.05-1.81)	.11
White European	853 (94.2)	4105 (94.6)			>28 (>84.8)	670 (95.6)		
Other ethnicities	53 (5.8)	236 (5.4)			<5 (<10)	14 (2)		
Highest qualification, n (%)			1.19 (1.11-1.29)	<.001			1.49 (1.15-1.98)	.004
None	29 (3.2)	301 (6.9)			<5 (<10)	164 (23.4)		
Any qualifications	861 (95)	2918 (67.2)			>28 (>84.8)	506 (72.2)		
Degree	412 (45.5)	1098 (25.3)			8 (24.2)	114 (16.3)		
Lifetime occupation, n (%)			1.04 (0.86-1.27)	.69			0.57 (0.2-1.47)	.26
Professional	344 (38)	985 (22.7)			5 (15.2)	70 (10)		
Other occupations	527 (58.2)	2167 (49.9)			28 (84.8)	559 (79.7)		
Never worked	29 (3.2)	84 (1.9)			0 (0)	53 (7.6)		
≥1 admission, n (%)	238 (26.2)	1298 (29.9)	1.05 (0.85-1.3)	.66	28 (84.8)	624 (89)	0.86 (0.27-3.88)	.82
Age of onset (years), median (IQR)[e]	17 (15)	19 (16)	0.99 (0.98-0.99)	.001	19 (13.5)	20 (11)	1 (0.95-1.05)	.99
Months since recruitment into parent study, median (IQR)[e]	38 (33)	48 (39)	0.98 (0.97-0.98)	<.001	108.5 (23.25)	110 (23.75)	1 (0.98-1.02)	.95

[a]Not all cells add up to the total N because of missing data.

[b]NCMH: National Centre for Mental Health.

[c]CoMPaSS: Cognition in Mood, Psychosis, and Schizophrenia Study.

[d]OR: odds ratio.

[e]Median and IQR are shown because of nonnormal distribution.

Table 4. Characteristics of participants and nonresponders (controls; N=939)[a].

Characteristic	NCMH[b] controls			
	Took part (n=213)	Did not take part (n=726)	OR[c] (95% CI)	*P* value
Age (years), mean (SD)	54.58 (17.24)	48.65 (17.69)	1.03 (1.01-1.04)	<.001
Women, n (%)	139 (65.3)	507 (69.8)	0.99 (0.65-1.52)	.97
Ethnicity, n (%)			1.19 (0.46-3.37)	.72
White European	205 (96.2)	673 (92.7)		
Other ethnicities	8 (3.8)	53 (7.3)		
Highest qualification, n (%)			1.5 (1.19-1.83)	<.001
None	<5 (<2.3)	7 (1)		
Any qualifications	>200 (>93.9)	302 (41.6)		
Degree	134 (62.9)	163 (22.5)		
Lifetime occupation, n (%)			0.71 (0.44-1.13)	.15
Professional	103 (48.4)	142 (19.6)		
Other occupations	99 (46.5)	167 (23)		
Never worked	6 (2.8)	<5 (<1)		
Months since recruitment into parent study, median (IQR)[d]	25 (10)	28 (13)	0.95 (0.93-0.96)	<.001

[a]Not all cells add up to the total N because of missing data.

[b]NCMH: National Centre for Mental Health.

[c]OR: odds ratio.

[d]Median and IQR are shown because of nonnormal distribution.

Completion Rates

Of the 961 participants who met the inclusion criteria, 887 (92.3%) completed at least one web-based cognitive task, and 644 (67%) completed all the tasks in the web-based battery. A breakdown of completion rates for each task and by diagnostic group can be found in Multimedia Appendix 1.

Cognitive Performance and Functioning

Linear regression indicated that lower cognitive performance (*g*) was associated with higher WHODAS 2.0 scores, indicating poorer functioning (B=−1.76, SE 0.26; *P*<.001). This association remained significant after covarying for diagnosis (B=−1.64, SE 0.26; *P*<.001). Higher scores on all tasks were associated with better functioning (Table 5).

Table 5. Association between cognitive performance and functioning (N=961).

Variable	Participants, n (%)	B[a] (SE)	*P* value
General cognitive performance (*g*)	561 (58.4)	−1.76 (0.26)	<.001
General cognitive performance (*g*) after covarying for diagnosis	561 (58.4)	−1.64 (0.26)	<.001
Individual tasks			
Digit symbol coding	645 (67.1)	−3.08 (0.47)	<.001
Verbal paired associates	557 (58)	−1.72 (0.50)	<.001
Backward digit span	528 (54.9)	−2.03 (0.51)	<.001
Hartshorne visual working memory test	499 (51.9)	−1.73 (0.51)	.008
Morphed emotion identification	565 (58.8)	−1.78 (0.53)	<.001
Matrix Reasoning Test	489 (50.9)	−1.54 (0.43)	<.001
Balloon Analogue Risk Task	487 (50.7)	−1.29 (0.39)	<.001
Multiple object tracking	459 (47.8)	−1.92 (0.43)	<.001
Vocabulary	470 (48.9)	−1.51 (0.51)	.003

[a]Regression coefficient.

Comparing Cognition Between Diagnostic Groups

There was a significant main effect of diagnosis for *g* ($F_{3,511}$=21.89; *P*<.001). Participants with depression had a lower performance than the controls (Hedges *g*=0.32; *P*=.01). The bipolar disorder group had lower performance than the controls (Hedges *g*=0.65; *P*<.001) and the depression group (Hedges *g*=0.33; *P*=.03). Participants with schizophrenia had the lowest performance relative to the controls (Hedges *g*=1.36; *P*<.001) and had lower performance than participants with depression (Hedges *g*=1.04; *P*<.001) and bipolar disorder (Hedges *g*=0.71; *P*=.002). Domain-specific effect sizes for pairwise comparisons are shown in Multimedia Appendix 1.

Discussion

Principal Findings

We developed a web-based cognitive battery for use in psychiatric research. The battery was designed to test the domains specified by the National Institute of Mental Health's MATRICS initiative. The aims of this study were to validate the web-based battery against the MCCB and evaluate whether web-based cognitive testing is suitable for research on psychiatric disorders. Our principal findings for each part of the study are outlined in the sections below.

Validation of the Web-Based Battery

We assessed convergent validity by conducting correlations between tasks that measured equivalent domains. Approximately 88% (7/8) of web-based tasks were correlated with the MCCB equivalent (*r*=0.26-0.73). This is comparable with the correlation coefficients reported in the validation of the UKB tasks (*r*=0.22-0.83) [16], web-based Cambridge Neuropsychological Test Automated Battery (*r*=0.39-0.73) [14], Amsterdam Cognition Scan (*r*=0.36-0.78) [15], and NutriCog Battery (*r*=0.42-0.73) [52]. However, only digit symbol coding, matrix reasoning, and vocabulary were most highly correlated with their equivalent offline tasks, and even these tasks were correlated with other MCCB tasks. The correlations between equivalent domains were also attenuated and, in some domains, close to 0 after correction for *g*. These findings suggest that at least some of the correlations between the web-based tasks and the MCCB were nonspecific and may reflect the tendency of cognitive tasks to be at least moderately correlated with a large proportion of the variance in performance accounted for by a higher-order factor (*g*) [53]. Overall, the battery may be better suited as a measure of general cognitive function *g* rather than a measure of specific cognitive domains.

There are two key differences between the web-based battery and the MCCB: (1) website versus offline administration and (2) use of different tasks to measure the same domain. Both differences are likely to affect the magnitude of the correlations between the tasks. Examining the correlation for speed of processing provides some insight into the extent to which the correlations are affected by these differences as the speed of processing domain was measured using offline and web-based versions of the same task, digit symbol coding. These tasks were the most highly correlated (*r*=0.73), suggesting that differences in web-based and traditional administration may not have had such a large impact on the magnitude of the correlations as differences arising from different tasks being used. This is consistent with the findings from the UKB and the Amsterdam Cognition Scan, showing that tasks with differing methodologies had lower correlations than those with similar methodologies [15,16]. A further consideration is the delay between administering the first and second batteries, as some participants did not complete the web-based battery on the second day as instructed. The correlations between the MCCB and the web-based battery did not change after controlling for the delay between the completion of the batteries, suggesting that the delay had little influence on the relationship between the tasks.

We observed the lowest nonspecific correlations between tasks that used different methodologies. For example, the morphed emotion identification task had a low correlation with the measure of social cognition in the MCCB, the Mayer–Salovey–Caruso Emotional Intelligence Test–Managing Emotions (MSCEIT-ME; *r*=0.26), but higher correlations with all other MCCB measures (*r*=0.33-0.56). Morphed emotion identification and MSCEIT-ME are designed to tap different aspects of social cognition. Morphed emotion identification is designed to measure a participant's ability to recognize emotional facial expressions [22], whereas the MSCEIT-ME was developed to measure emotion self-regulation [54]. We did not select a formally equivalent web-based task for the MSCEIT-ME for 2 reasons. First, although previous research has identified impaired performance on the MSCEIT-ME in participants with schizophrenia [19,55], studies have not identified impairments in this task for participants with bipolar disorder [19,56,57]. Therefore, we did not consider it a suitable measure for this web-based battery, which was designed to measure cognitive function in participants with a range of psychiatric disorders. Second, in our experience of administering the MSCEIT-ME to >1000 participants with psychosis, we found that participants frequently required guidance and explanations of the scenarios, which would not be practical for a web-based, unsupervised measure. Studies have identified impairments in emotion recognition in participants with bipolar disorder [58,59], depression [60], autism [61], and posttraumatic stress disorder [62]. Therefore, we considered this a more suitable measure for cross-disorder research. However, we did not find evidence of impairments in this task for participants with depression or bipolar disorder in this study as effect sizes were small (*d*=0.17 for depression; *d*=0.28 for bipolar disorder) and nonsignificant (see Multimedia Appendix 1).

The Hartshorne visual working memory task was not correlated with its selected equivalent in the MCCB, the Brief Visuospatial Memory Test–Revised (BVMT-R). The BVMT-R is an immediate visual recall task in which participants are presented with shapes for 10 seconds and asked to draw these shapes from memory [63]. Drawings are rated based on the accuracy of recall and memory of the location of the shapes. It was not possible to select a task that would entirely replicate the BVMT-R because of the difficulties in automating the study administration and scoring on the web and the possibility that participants may cheat in an unsupervised setting by copying the shapes while they are being displayed onscreen. Therefore, the Hartshorne

visual working memory task was chosen as an alternative as performance on this task also relies on short-term memory of both shapes and locations. However, the Hartshorne visual working memory task is more complex as, in addition to memorizing shapes and locations, the participant must identify whether a new target shape is the same or different from the shape that previously occupied that position [23]. As such, the task incorporates aspects of working memory and problem solving. Our results suggest that these tasks are not comparable and that the Hartshorne visual working memory task is not a suitable alternative measure of immediate visual recall.

The BART was not correlated with any of the MCCB tasks. This was not surprising, given that the BART is a behavioral measure rather than a neurocognitive task, and the MCCB is primarily made up of neurocognitive measures. Nevertheless, the BART may be a useful measure of risk-taking behavior that does not rely on self-report, as it has been shown to be correlated with measures of sensation seeking, impulsivity, behavior constraint, and actual risk-taking behaviors [24].

Feasibility of Web-Based Cognitive Testing

We demonstrated that web-based cognitive testing is an effective method of collecting data from a large sample of participants with psychiatric disorders. To date, we have obtained cognitive data from >1000 participants diagnosed with a range of psychiatric disorders. A key concern with web-based testing in psychiatric research is whether the sample is representative, particularly given evidence that there is a digital divide between patients with psychiatric disorders and the wider UK population [64]. Across all samples (CoMPaSS and NCMH cases and controls), a higher level of education was associated with web-based participation. This finding suggests that earlier concerns about education bias in web-based samples may be correct and is a clear limitation of web-based testing. This issue is not exclusive to web-based cognitive testing, as similar recruitment biases have been identified in mental health studies using web-based questionnaires [17], and response rates to research invitations have been shown to be positively associated with educational attainment more generally [65]. The web-based control group was older and had been recruited more recently into the NCMH than nonresponders. Finally, participants with psychiatric disorders recruited from the NCMH were older, more likely to be female, less likely to be White European, and had been recruited more recently than nonresponders. The OR for age was close to 1, which does not support early preconceptions that internet samples would be overrepresented by younger people [11]. Our web-based sample was drawn from 2 existing clinical studies, and it should be noted that these original studies also have recruitment bias, although there is some evidence that the CoMPaSS sample is representative of the wider population of patients with psychosis in Wales based on the linkage of these data with routinely collected records [66]. Nevertheless, we were unable to assess whether the participants in our web-based study were representative of the wider population of patients with psychiatric disorders. However, our results do indicate that web-based samples may have recruitment bias beyond that seen in traditional clinical studies of psychiatric disorders.

In terms of clinical differences between participants and nonresponders, we did not find evidence of differences in hospital admissions. The difference in age of onset was significant; however, the OR was close to 1, indicating that the difference was very small. There were differences in the proportion of diagnoses, as individuals with schizophrenia were less likely to participate, and individuals with an eating disorder or comorbid depression and anxiety were more likely to participate. The response rates for individuals with other psychotic disorders were also low. These response rates may reflect the severity of illness. We did not find differences in hospital admissions as a binary measure but were unable to examine the number of admissions, length of hospitalizations, or whether admissions were under the Mental Health Act (as data were only available for a small proportion of the sample), which would have provided more detailed information on the severity of illness. Although this is another limitation of web-based testing in a mental health sample, an advantage of web-based testing is that the tests can be administered anywhere with internet access, supervised or unsupervised. Sample representativeness may be improved by providing opportunities for participants to take part in a supervised setting, such as a psychiatric clinic or research facility, if they lack the skills or resources to access the internet unsupervised. A combination of approaches (supervised and unsupervised and clinical or home settings) may reduce the financial and logistical burden of assessing cognitive function in large cohorts associated with traditional studies while limiting the recruitment bias associated with purely web-based studies. Studies using web-based methods for data collection should consider providing additional support to individuals with more severe mental illnesses, such as psychosis.

The completion rates give an indication of the tolerability of the cognitive battery. Of the 961 eligible participants who consented to the study, 887 (92.3%) completed at least one task, which is similar to the 87% reported by another web-based cognitive study [67]. Most participants completed all 9 tasks in the cognitive battery (644/961, 67% of eligible participants or 644/887, 72.6% of participants who started the tasks). This figure is comparable with the completion rates reported by the Twins Early Development Study for their web-based battery of 8 tasks (65%), although the study assessed children [9]. This figure is lower than those reported in face-to-face cognitive studies of participants with psychiatric disorders [68,69]. A lower completion rate was expected, given that the participants were unsupervised and would not have the support of a researcher to complete the study. However, this should be considered a potential limitation of web-based testing in psychiatric research, and more work is needed to understand who is likely to drop out; the reasons for dropout; and whether any measures can be taken to mitigate dropout; for example, by reducing the overall length of the battery. Overall, the completion rates were adequate, suggesting that the tasks were well-tolerated by most participants. All participants who did not complete the tasks were followed up by email or phone, and technical issues were recorded and resolved where possible. The number of technical issues reported by participants was small (part 1: 6 issues reported; part 2: 9 issues reported). In total, 19 participants reported technical problems, of which 13

(68%) were able to complete the battery (part 1: 6, 32% reported [9% of validation sample] and 5, 26% completed; part 2: 13, 68% reported [1.5% of sample] and 8, 42% completed).

Lower cognitive performance was associated with poorer functioning. These results suggest that performance on the battery is an important indicator of overall functioning. The results are consistent with a prospective study that reported an association between cognitive performance and WHODAS 2.0 scores in a cross-disorder sample of participants with depression, bipolar disorder, and psychosis [70]. In this study, performance on digit symbol coding and backward digit span was most strongly associated with scores on the WHODAS 2.0. Both digit symbol coding and backward digit span have short administration times and thus may be particularly suited for brief assessments of cognition in a clinical setting. It should be noted that this study was cross-sectional; thus, it was not possible to examine whether changes in performance on the battery were associated with changes in functioning.

We compared cognitive performance in participants with major depressive disorder, bipolar disorder, and schizophrenia as these groups have been extensively assessed for research using traditional cognitive assessments. Our results demonstrated a pattern of decreasing cognitive scores from major depressive disorder to bipolar disorder to schizophrenia and are consistent with studies showing lower performance in schizophrenia compared with major depressive disorder and bipolar disorder [70,71]. This suggests that the performance of the battery can be used to discriminate cases and controls. However, this pattern was not consistent in the analyses of individual domains. Effect sizes for the depression group across domains (Hedges g=0.07-0.39) were also lower than those reported by previous meta-analyses (effect sizes ranged from 0.32 to 0.97 across domains [3,72,73]), which may be explained by recruitment bias or the self-report nature of the NCMH diagnoses.

Limitations

In addition to the limitations discussed above, several further limitations should be noted. Test–retest reliability of the web-based battery was not assessed in this study. In assessing validity, the correlation between a new task and the gold standard task cannot exceed √(reliability of reference task×reliability of new task) [10]. Therefore, the upper limits of the correlations between tasks were unknown. This is helpful for interpretation but does not change the magnitude or significance of the correlations. The order of completion of the batteries was not counterbalanced in the validation study for practical reasons, which may mean that performance on the web-based battery was subject to practice effects, particularly those tasks with similar methodologies. However, we would expect practice effects to be minimal as most web-based and MCCB tasks used different methodologies, and none of the tasks used the same stimuli. In the NCMH sample, diagnosis was based on self-report rather than structured interviews, which may result in incomplete or inaccurate diagnoses. However, participants were asked to report diagnoses that they had been given by a health professional, which is consistent with the approach taken by other large studies with self-report measures of diagnosis, such as the UKB [25,26]. There was a smaller schizophrenia group because of a lower response rate from participants with this disorder. A further limitation of web-based testing is the unsupervised environment in which the data are collected, which makes it difficult to minimize distractions or cheating. A study by Germine et al [12] using the same platform (TestMyBrain) found low levels of self-reported cheating, and their data were not consistent with widespread cheating. We selected tasks that would minimize the possibility of cheating where possible and also instructed participants to complete the tasks in a quiet environment. In addition, most participants completed both batteries at home to minimize the differences in the test settings.

Conclusions

The web-based battery has several strengths, including the use of tasks taken from published research, domains selected based on the MATRICS initiative, and compatibility with a range of devices, including touchscreen devices. The availability of demographic and clinical data on individuals who did not participate in the study was a unique strength that allowed us to assess potential recruitment bias.

In conclusion, we developed a new web-based cognitive battery and used it to collect data from a large sample of participants with a range of psychiatric disorders. There was some evidence of recruitment bias, and the levels of impairment found in the depression group were less severe than those reported by traditional face-to-face studies. However, the battery provided a reliable measure of g, completion rates were adequate, our findings were consistent with studies using traditional assessments, and feedback from participants was positive. Our findings offer valuable insights into the advantages and disadvantages of testing cognitive function remotely for mental health research, which is particularly important given the increasing number of psychiatric studies using digital methods of assessment. In the next stage of development, we intend to use these findings to reduce the size of the battery to a briefer version removing tasks with low correlations, rated poorly by participants, or those presenting technical issues. Given that the correlations between some web-based tasks and the MCCB were nonspecific, the battery will include the tasks that were best suited for measuring their domain and providing a measure of g. We will also redesign the battery with a user-friendly interface with input from patient representative groups.

Cognitive impairments are one of the causes of long-term disability among patients with psychiatric disorders, particularly schizophrenia [7,8]. In the United Kingdom, assessment of cognitive skills has been recommended in clinical settings for patients with psychosis [74]. However, there are barriers associated with accessing appropriate cognitive assessments, such as cost, licensing, and the lack of available assessments that provide clinically meaningful feedback on performance [75]. It has been suggested that web-based tools may be a cost-effective solution [75]. Our results indicate that this assessment is suitable for detecting impairments in individuals with schizophrenia. Therefore, it is our intention to explore the potential clinical utility of the battery in future work.

Acknowledgments

This work was supported by a Medical Research Council (MRC) PhD studentship to AJL and by the National Centre for Mental Health (NCMH). The NCMH is a collaboration between the Cardiff, Swansea, and Bangor Universities and is funded by the Welsh Government through Health and Care Research Wales. The NCMH provided resources and staff to assist with the development of the battery and recruitment of participants for both stages of the study. The further development of the battery was supported by 2 MRC grants, the MRC Mental Health Data Pathfinder award (MC_PC_17212) and the MRC Confidence in Concept award (MC_PC_17170). The MRC Pathfinder award (MC_PC_17212) funded AJL's postdoctoral position. The MRC Confidence in Concept award (MC_PC_17170) is a proof-of-concept grant, which funded the further development of the cognitive assessment, including creating a user-friendly interface and supporting public involvement. The work at Cardiff University was funded by the MRC Centre (MR/L010305/1) and program grant (G0500509). This research was conducted with the support of the web-based platform TestMyBrain, a resource maintained and supported by The Many Brains Project, Inc. The web-based tasks remain the intellectual property of The Many Brains Project, which retains the copyright for these tasks and the source code. The authors thank Dr Laura Germine for her support in developing this battery. The authors thank the Cognition in Mood, Psychosis, and Schizophrenia Study and NCMH study participants for their invaluable contribution to this project. The authors also thank the NCMH field team, Dr Catrin Lewis, Dr Lawrence Raisanen, and Karin Alderson, for their support in recruiting participants for this study.

Authors' Contributions

AJL is the lead author and was involved in all aspects of the study, including designing the assessment tools and study methodology, overseeing the recruitment of participants, conducting analysis and interpretation of the data, and drafting and redrafting the manuscript. JTRW is the senior author; the principal investigator of the Cognition in Mood, Psychosis, and Schizophrenia Study; deputy director of the National Centre for Mental Health; and was involved in all aspects of the paper. IRJ is the chief investigator and director of the National Centre for Mental Health and advised on the methodology and interpretation of the results. All authors critically revised the paper and approved the final version to be submitted.

Conflicts of Interest

IRJ and JTRW have received grant funding from Takeda Pharmaceuticals for research unrelated to this manuscript.

Multimedia Appendix 1

Detailed information about the parent studies, recruitment response rates, inclusion criteria, tasks included in the web-based battery, completion rates, and supplementary results.

[DOCX File , 92 KB - jmir_v24i2e28233_app1.docx]

References

1. Heinrichs RW, Zakzanis KK. Neurocognitive deficit in schizophrenia: a quantitative review of the evidence. Neuropsychology 1998;12(3):426-445. [doi: 10.1037/0894-4105.12.3.426]
2. Bora E, Yucel M, Pantelis C. Cognitive endophenotypes of bipolar disorder: a meta-analysis of neuropsychological deficits in euthymic patients and their first-degree relatives. J Affect Disord 2009 Feb;113(1-2):1-20. [doi: 10.1016/j.jad.2008.06.009] [Medline: 18684514]
3. Rock PL, Roiser JP, Riedel WJ, Blackwell AD. Cognitive impairment in depression: a systematic review and meta-analysis. Psychol Med 2013 Oct 29;44(10):2029-2040. [doi: 10.1017/s0033291713002535]
4. Keefe RS, Malhotra AK, Meltzer HY, Kane JM, Buchanan RW, Murthy A, et al. Efficacy and safety of donepezil in patients with schizophrenia or schizoaffective disorder: significant placebo/practice effects in a 12-week, randomized, double-blind, placebo-controlled trial. Neuropsychopharmacology 2008 May 11;33(6):1217-1228. [doi: 10.1038/sj.npp.1301499] [Medline: 17625502]
5. Keefe RS, Bilder RM, Davis SM, Harvey PD, Palmer BW, Gold JM, CATIE Investigators, Neurocognitive Working Group. Neurocognitive effects of antipsychotic medications in patients with chronic schizophrenia in the CATIE Trial. Arch Gen Psychiatry 2007 Jun 01;64(6):633-647. [doi: 10.1001/archpsyc.64.6.633] [Medline: 17548746]
6. Goldberg JK, Chengappa KN. Identifying and treating cognitive impairment in bipolar disorder. Bipolar Disord 2009 Jun;11 Suppl 2:123-137. [doi: 10.1111/j.1399-5618.2009.00716.x] [Medline: 19538691]
7. Depp C, Mausbach B, Harmell A, Savla G, Bowie C, Harvey P, et al. Meta-analysis of the association between cognitive abilities and everyday functioning in bipolar disorder. Bipolar Disord 2012 May;14(3):217-226 [FREE Full text] [doi: 10.1111/j.1399-5618.2012.01011.x] [Medline: 22548895]
8. Fett AJ, Viechtbauer W, Dominguez M, Penn DL, van Os J, Krabbendam L. The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: a meta-analysis. Neurosci Biobehav Rev 2011 Jan;35(3):573-588. [doi: 10.1016/j.neubiorev.2010.07.001] [Medline: 20620163]

9. Haworth CM, Harlaar N, Kovas Y, Davis OS, Oliver BR, Hayiou-Thomas ME, et al. Internet cognitive testing of large samples needed in genetic research. Twin Res Hum Genet 2007 Aug 21;10(4):554-563. [doi: 10.1375/twin.10.4.554] [Medline: 17708696]
10. Feenstra HE, Vermeulen IE, Murre JM, Schagen SB. Online cognition: factors facilitating reliable online neuropsychological test results. Clin Neuropsychol 2017 Jan 08;31(1):59-84. [doi: 10.1080/13854046.2016.1190405] [Medline: 27266677]
11. Gosling SD, Vazire S, Srivastava S, John OP. Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. Am Psychol 2004;59(2):93-104. [doi: 10.1037/0003-066x.59.2.93]
12. Germine L, Nakayama K, Duchaine BC, Chabris CF, Chatterjee G, Wilmer JB. Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. Psychon Bull Rev 2012 Oct 25;19(5):847-857. [doi: 10.3758/s13423-012-0296-9] [Medline: 22829343]
13. Darby D, Fredrickson J, Pietrzak R, Maruff P, Woodward M, Brodtmann A. Reliability and usability of an internet-based computerized cognitive testing battery in community-dwelling older people. Comput Hum Behav 2014 Jan;30:199-205. [doi: 10.1016/j.chb.2013.08.009]
14. Backx R, Skirrow C, Dente P, Barnett JH, Cormack FK. Comparing web-based and lab-based cognitive assessment using the Cambridge neuropsychological test automated battery: a within-subjects counterbalanced study. J Med Internet Res 2020 Aug 04;22(8):e16792 [FREE Full text] [doi: 10.2196/16792] [Medline: 32749999]
15. Feenstra HE, Murre JM, Vermeulen IE, Kieffer JM, Schagen SB. Reliability and validity of a self-administered tool for online neuropsychological testing: the Amsterdam Cognition Scan. J Clin Exp Neuropsychol 2018 Apr 03;40(3):253-273. [doi: 10.1080/13803395.2017.1339017] [Medline: 28671504]
16. Fawns-Ritchie C, Deary IJ. Reliability and validity of the UK Biobank cognitive tests. PLoS One 2020 Apr 20;15(4):e0231627 [FREE Full text] [doi: 10.1371/journal.pone.0231627] [Medline: 32310977]
17. Batterham PJ. Recruitment of mental health survey participants using internet advertising: content, characteristics and cost effectiveness. Int J Methods Psychiatr Res 2014 Jun 24;23(2):184-191 [FREE Full text] [doi: 10.1002/mpr.1421] [Medline: 24615785]
18. Underwood JF, Kendall KM, Berrett J, Lewis C, Anney R, van den Bree MB, et al. Autism spectrum disorder diagnosis in adults: phenotype and genotype findings from a clinically derived cohort. Br J Psychiatry 2019 Nov 26;215(5):647-653 [FREE Full text] [doi: 10.1192/bjp.2019.30] [Medline: 30806336]
19. Lynham AJ, Hubbard L, Tansey KE, Hamshere ML, Legge SE, Owen MJ, et al. Examining cognition across the bipolar/schizophrenia diagnostic spectrum. J Psychiatry Neurosci 2018 Jul 01;43(4):245-253. [doi: 10.1503/jpn.170076]
20. Wing JK, Babor T, Brugha T, Burke J, Cooper JE, Giel R, et al. SCAN. Schedules for clinical assessment in neuropsychiatry. Arch Gen Psychiatry 1990 Jun 01;47(6):589-593. [doi: 10.1001/archpsyc.1990.01810180089012] [Medline: 2190539]
21. Nelson HE, Willison J. Nart: National Adult Reading Test. Windsor: NFER - Nelson; 1991.
22. Kessels RP, Montagne B, Hendriks AW, Perrett DI, de Haan EH. Assessment of perception of morphed facial expressions using the Emotion Recognition Task: normative data from healthy participants aged 8-75. J Neuropsychol 2014 Mar 14;8(1):75-93. [doi: 10.1111/jnp.12009] [Medline: 23409767]
23. Hartshorne JK. Visual working memory capacity and proactive interference. PLoS One 2008 Jul 23;3(7):e2716 [FREE Full text] [doi: 10.1371/journal.pone.0002716] [Medline: 18648493]
24. Lejuez CW, Read JP, Kahler CW, Richards JB, Ramsey SE, Stuart GL, et al. Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART). J Experiment Psychol Appl 2002;8(2):75-84. [doi: 10.1037/1076-898x.8.2.75]
25. Protocol for a large-scale prospective epidemiological resource - Protocol No: UKBB-PROT-09-06 (Main Phase). UK Biobank. 2007. URL: https://www.ukbiobank.ac.uk/media/gnkeyh2q/study-rationale.pdf [accessed 2021-12-29]
26. Verbal interview stage - version 1.1. UK Biobank. 2012. URL: https://biobank.ctsu.ox.ac.uk/crystal/ukb/docs/Interview.pdf [accessed 2021-12-29]
27. Rees E, Walters JT, Georgieva L, Isles AR, Chambert KD, Richards AL, et al. Analysis of copy number variations at 15 schizophrenia-associated loci. Br J Psychiatry 2014 Feb 02;204(2):108-114 [FREE Full text] [doi: 10.1192/bjp.bp.113.131052] [Medline: 24311552]
28. Passell E, Dillon D, Baker J, Vogel S, Scheuer L, Mirin N, et al. Digital cognitive assessment: results from the testmybrain nimh research domain criteria (rdoc) field test battery report. PsyArXiv. 2019. URL: https://psyarxiv.com/dcszr/ [accessed 2021-12-29]
29. Joy S, Kaplan E, Fein D. Speed and memory in the WAIS-III Digit Symbol--Coding subtest across the adult lifespan. Arch Clin Neuropsychol 2004 Sep;19(6):759-767. [doi: 10.1016/j.acn.2003.09.009] [Medline: 15288329]
30. Frigerio E, Burt D, Montagne B, Murray LK, Perrett DI. Facial affect perception in alcoholics. Psychiatry Res 2002 Dec;113(1-2):161-171. [doi: 10.1016/s0165-1781(02)00244-5]
31. Nuechterlein KH, Barch DM, Gold JM, Goldberg TE, Green MF, Heaton RK. Identification of separable cognitive factors in schizophrenia. Schizophr Res 2004 Dec 15;72(1):29-39. [doi: 10.1016/j.schres.2004.09.007] [Medline: 15531405]
32. Wechsler D. WAIS-III: Administration and Scoring Manual : Wechsler Adult Intelligence Scale. San Antonio, TX, United States: Psychological Corporation; 1997.

33. Wechsler DC. WASI II: Wechsler Abbreviated Scale of Intelligence. 2nd Ed. San Antonio, TX, United States: Psychological Corporation; 2011.
34. Lejuez C, Aklin WM, Zvolensky MJ, Pedulla CM. Evaluation of the Balloon Analogue Risk Task (BART) as a predictor of adolescent real-world risk-taking behaviours. J Adolesc 2003 Aug;26(4):475-479. [doi: 10.1016/s0140-1971(03)00036-8]
35. Wilmer J, Martini P, Germine L, Nakayama K. Multiple object tracking predicts math potential. J Vision 2016 Aug 31;16(12):421. [doi: 10.1167/16.12.421]
36. Bright P, Jaldow E, Kopelman MD. The National Adult Reading Test as a measure of premorbid intelligence: a comparison with estimates derived from demographic variables. J Int Neuropsychol Soc 2002 Sep 12;8(6):847-854. [doi: 10.1017/s1355617702860131] [Medline: 12240749]
37. Russell A, Munro J, Jones P, Hayward P, Hemsley D, Murray R. The National Adult Reading Test as a measure of premorbid IQ in schizophrenia. Br J Clin Psychol 2000 Sep;39(3):297-305. [doi: 10.1348/014466500163301] [Medline: 11033751]
38. Hartshorne JK, Germine LT. When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. Psychol Sci 2015 Apr 13;26(4):433-443 [FREE Full text] [doi: 10.1177/0956797614567339] [Medline: 25770099]
39. Malhotra N, Krosnick J, Haertel E. The psychometric properties of the gss wordsum vocabulary test. GSS Methodological Report. 2007. URL: http://gss.norc.org/Documents/reports/methodological-reports/MR111%20The%20Psychometric%20Properties%20of%20the%20GSS%20Wordsum%20Vocabulary%20Test.pdf [accessed 2021-12-29]
40. The Many Brains Project. Testmybrain. URL: https://www.testmybrain.org [accessed 2020-10-28]
41. Marder SR, Fenton W. Measurement and treatment research to improve cognition in schizophrenia: NIMH MATRICS initiative to support the development of agents for improving cognition in schizophrenia. Schizophr Res 2004 Dec 15;72(1):5-9. [doi: 10.1016/j.schres.2004.09.010] [Medline: 15531402]
42. Nuechterlein KH, Green MF, Kern RS, Baade LE, Barch DM, Cohen JD, et al. The MATRICS consensus cognitive battery, part 1: test selection, reliability, and validity. Am J Psychiatry 2008 Feb;165(2):203-213. [doi: 10.1176/appi.ajp.2007.07010042] [Medline: 18172019]
43. Rehm J, Üstün TB, Saxena S, Nelson CB, Chatterji S, Ivis F, et al. On the development and psychometric testing of the WHO screening instrument to assess disablement in the general population. Int J Method Psychiat Res 2006 Jun;8(2):110-122. [doi: 10.1002/mpr.61]
44. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983 Jun;67(6):361-370. [doi: 10.1111/j.1600-0447.1983.tb09716.x] [Medline: 6880820]
45. Altman EG, Hedeker D, Peterson JL, Davis JM. The Altman Self-Rating Mania Scale. Biol Psychiatry 1997 Nov;42(10):948-955. [doi: 10.1016/s0006-3223(96)00548-3]
46. Kendall KM, Rees E, Bracher-Smith M, Legge S, Riglin L, Zammit S, et al. Association of rare copy number variants with risk of depression. JAMA Psychiatry 2019 Aug 01;76(8):818-825 [FREE Full text] [doi: 10.1001/jamapsychiatry.2019.0566] [Medline: 30994872]
47. Richards AL, Pardiñas AF, Frizzati A, Tansey KE, Lynham AJ, Holmans P, GROUP Investigators, EUGEI WP2 Group, Schizophrenia Working Group of the Psychiatric Genomics Consortium, et al. The relationship between polygenic risk scores and cognition in schizophrenia. Schizophr Bull 2020 Feb 26;46(2):336-344 [FREE Full text] [doi: 10.1093/schbul/sbz061] [Medline: 31206164]
48. Tucker-Drob EM, Salthouse TA. Confirmatory factor analysis and multidimensional scaling for construct validation of cognitive abilities. Int J Behav Dev 2009 Feb 25;33(3):277-285 [FREE Full text] [doi: 10.1177/0165025409104489] [Medline: 20963182]
49. Lu ZQ. The elements of statistical learning: data mining, inference, and prediction. J Royal Stat Society Ser A 2010;173(3):693-694. [doi: 10.1111/j.1467-985X.2010.00646_6.x]
50. Hedges LI. Statistical Methods for Meta-Analysis. Orlando, FL: Academic Press; 1985.
51. MRC Centre for Neuropsychiatric Genetics and Genomics. Walters Group Data Repository. URL: https://walters.psycm.cf.ac.uk/ [accessed 2022-01-11]
52. Assmann KE, Bailet M, Lecoffre AC, Galan P, Hercberg S, Amieva H, et al. Comparison between a self-administered and supervised version of a web-based cognitive test battery: results from the NutriNet-Santé Cohort Study. J Med Internet Res 2016 Apr 05;18(4):e68 [FREE Full text] [doi: 10.2196/jmir.4862] [Medline: 27049114]
53. Deary IJ. Human intelligence differences: a recent history. Trends Cogn Sci 2001 Mar;5(3):127-130. [doi: 10.1016/s1364-6613(00)01621-1]
54. DeTore NR, Mueser KT, McGurk SR. What does the Managing Emotions branch of the MSCEIT add to the MATRICS consensus cognitive battery? Schizophr Res 2018 Jul;197:414-420. [doi: 10.1016/j.schres.2018.02.018] [Medline: 29486955]
55. Kern RS, Gold JM, Dickinson D, Green MF, Nuechterlein KH, Baade LE, et al. The MCCB impairment profile for schizophrenia outpatients: results from the MATRICS psychometric and standardization study. Schizophr Res 2011 Mar;126(1-3):124-131 [FREE Full text] [doi: 10.1016/j.schres.2010.11.008] [Medline: 21159492]

56. Burdick KE, Goldberg TE, Cornblatt BA, Keefe RS, Gopin CB, Derosse P, et al. The MATRICS consensus cognitive battery in patients with bipolar I disorder. Neuropsychopharmacology 2011 Jul 30;36(8):1587-1592 [FREE Full text] [doi: 10.1038/npp.2011.36] [Medline: 21451499]
57. Van Rheenen TE, Rossell SL. An empirical evaluation of the MATRICS Consensus Cognitive Battery in bipolar disorder. Bipolar Disord 2014 May 09;16(3):318-325. [doi: 10.1111/bdi.12134] [Medline: 24119238]
58. Samamé C, Martino D, Strejilevich S. Social cognition in euthymic bipolar disorder: systematic review and meta-analytic approach. Acta Psychiatr Scand 2012 Apr;125(4):266-280. [doi: 10.1111/j.1600-0447.2011.01808.x] [Medline: 22211280]
59. Addington J, Addington D. Facial affect recognition and information processing in schizophrenia and bipolar disorder. Schizophrenia Res 1998 Aug;32(3):171-181. [doi: 10.1016/s0920-9964(98)00042-5]
60. Dalili MN, Penton-Voak IS, Harmer CJ, Munafò MR. Meta-analysis of emotion recognition deficits in major depressive disorder. Psychol Med 2014 Nov 14;45(6):1135-1144. [doi: 10.1017/s0033291714002591]
61. Lozier LM, Vanmeter JW, Marsh AA. Impairments in facial affect recognition associated with autism spectrum disorders: a meta-analysis. Dev Psychopathol 2014 Jun 10;26(4pt1):933-945. [doi: 10.1017/s0954579414000479]
62. Plana I, Lavoie M, Battaglia M, Achim AM. A meta-analysis and scoping review of social cognition performance in social phobia, posttraumatic stress disorder and other anxiety disorders. J Anxiety Disord 2014 Mar;28(2):169-177. [doi: 10.1016/j.janxdis.2013.09.005] [Medline: 24239443]
63. Benedict RH, Schretlen D, Groninger L, Dobraski M, Shpritz B. Revision of the brief visuospatial memory test: studies of normal performance, reliability, and validity. Psychol Assess 1996;8(2):145-153. [doi: 10.1037/1040-3590.8.2.145]
64. Robotham D, Satkunanathan S, Doughty L, Wykes T. Do we still have a digital divide in mental health? A five-year survey follow-up. J Med Internet Res 2016 Nov 22;18(11):e309 [FREE Full text] [doi: 10.2196/jmir.6511] [Medline: 27876684]
65. Patel MX, Doku V, Tennakoon L. Challenges in recruitment of research participants. Adv Psychiatr Treat 2018 Jan 02;9(3):229-238. [doi: 10.1192/apt.9.3.229]
66. Lloyd K, McGregor J, John A, Craddock N, Walters JT, Linden D, et al. A national population-based e-cohort of people with psychosis (PsyCymru) linking prospectively ascertained phenotypically rich and genetic data to routinely collected records: overview, recruitment and linkage. Schizophr Res 2015 Aug;166(1-3):131-136. [doi: 10.1016/j.schres.2015.05.036] [Medline: 26139629]
67. Lumsden J, Skinner A, Coyle D, Lawrence N, Munafo M. Attrition from web-based cognitive testing: a repeated measures comparison of gamification techniques. J Med Internet Res 2017 Nov 22;19(11):e395 [FREE Full text] [doi: 10.2196/jmir.8473] [Medline: 29167090]
68. Keefe RS, Fox KH, Harvey PD, Cucchiaro J, Siu C, Loebel A. Characteristics of the MATRICS Consensus Cognitive Battery in a 29-site antipsychotic schizophrenia clinical trial. Schizophr Res 2011 Feb;125(2-3):161-168. [doi: 10.1016/j.schres.2010.09.015] [Medline: 21075600]
69. Keefe R, Goldberg TE, Harvey PD, Gold JM, Poe MP, Coughenour L. The brief assessment of cognition in schizophrenia: reliability, sensitivity, and comparison with a standard neurocognitive battery. Schizophr Res 2004 Jun 01;68(2-3):283-297. [doi: 10.1016/j.schres.2003.09.011] [Medline: 15099610]
70. Lee RS, Hermens DF, Naismith SL, Lagopoulos J, Jones A, Scott J, et al. Neuropsychological and functional outcomes in recent-onset major depression, bipolar disorder and schizophrenia-spectrum disorders: a longitudinal cohort study. Transl Psychiatry 2015 Apr 28;5(4):e555 [FREE Full text] [doi: 10.1038/tp.2015.50] [Medline: 25918992]
71. Hill SK, Reilly JL, Harris MS, Rosen C, Marvin RW, Deleon O, et al. A comparison of neuropsychological dysfunction in first-episode psychosis patients with unipolar depression, bipolar disorder, and schizophrenia. Schizophr Res 2009 Sep;113(2-3):167-175 [FREE Full text] [doi: 10.1016/j.schres.2009.04.020] [Medline: 19450952]
72. Bora E, Harrison BJ, Yücel M, Pantelis C. Cognitive impairment in euthymic major depressive disorder: a meta-analysis. Psychol Med 2012 Oct 26;43(10):2017-2026. [doi: 10.1017/s0033291712002085]
73. Snyder HR. Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: a meta-analysis and review. Psychol Bull 2013 Jan;139(1):81-132 [FREE Full text] [doi: 10.1037/a0028727] [Medline: 22642228]
74. Psychosis and schizophrenia in adults: prevention and management. National Institute for Health and Care Excellence. 2014. URL: https://www.nice.org.uk/guidance/cg178/resources/psychosis-and-schizophrenia-in-adults-prevention-and-management-pdf-35109758952133 [accessed 2021-05-24]
75. Keefe RS, Haig GM, Marder SR, Harvey PD, Dunayevich E, Medalia A, et al. Report on ISCTM Consensus Meeting on clinical assessment of response to treatment of cognitive impairment in schizophrenia. Schizophr Bull 2016 Jan 11;42(1):19-33 [FREE Full text] [doi: 10.1093/schbul/sbv111] [Medline: 26362273]

Abbreviations

BART: Balloon Analogue Risk Task
BVMT-R: Brief Visuospatial Memory Test–Revised
CoMPaSS: Cognition in Mood, Psychosis, and Schizophrenia Study
MATRICS: Measurement and Treatment Research to Improve Cognition in Schizophrenia

MCCB: Measurement and Treatment Research to Improve Cognition in Schizophrenia Consensus Cognitive Battery
MDS: multidimensional scaling
MRC: Medical Research Council
MSCEIT-ME: Mayer–Salovey–Caruso Emotional Intelligence Test–Managing Emotions
NART: National Adult Reading Test
NCMH: National Centre for Mental Health
OR: odds ratio
UKB: UK Biobank
WHODAS 2.0: World Health Organization Disability Assessment Schedule Version 2

Edited by R Kukafka; submitted 26.02.21; peer-reviewed by A Wright, C Jones; comments to author 11.05.21; revised version received 11.06.21; accepted 21.11.21; published 10.02.22.

Please cite as:
Lynham AJ, Jones IR, Walters JTR
Web-Based Cognitive Testing in Psychiatric Research: Validation and Usability Study
J Med Internet Res 2022;24(2):e28233
URL: https://www.jmir.org/2022/2/e28233
doi:10.2196/28233
PMID:35142640

Original Paper

Internet-Based Audiologist-Guided Cognitive Behavioral Therapy for Tinnitus: Randomized Controlled Trial

Eldré W Beukes[1,2], BSc, MSc, PhD; Gerhard Andersson[3], BSc, MSc, PhD; Marc Fagelson[4,5], BSc, MSc, PhD; Vinaya Manchaiah[2,6,7,8,9], BSc, MSc, PhD

[1]Vision and Hearing Research Centre, Anglia Ruskin University, Cambridge, United Kingdom

[2]Virtual Hearing Lab, Collaborative Initiative Between University of Colorado School of Medicine and University of Pretoria, Aurora, CO, United States

[3]Department of Behavioral Sciences and Learning, Linköping University, Linköping, Sweden

[4]Department of Audiology and Speech-Language Pathology, East Tennessee State University, Johnson City, TN, United States

[5]Auditory Vestibular Research Enhancement Award Program, Audiological Rehabilitation Laboratory, Veterans Affairs Medical Center, Mountain Home, TN, United States

[6]Department of Speech-Language Pathology and Audiology, University of Pretoria, Pretoria, South Africa

[7]Department of Otolaryngology–Head and Neck Surgery, University of Colorado School of Medicine, Aurora, CO, United States

[8]UCHealth Hearing and Balance, University of Colorado Hospital, Aurora, CO, United States

[9]Department of Speech and Hearing, Manipal College of Health Professions, Manipal Academy of Higher Education, Manipal, India

Corresponding Author:
Eldré W Beukes, BSc, MSc, PhD
Vision and Hearing Research Centre
Anglia Ruskin University
East Road
Cambridge, CB1 1TP
United Kingdom
Phone: 44 07951113157
Email: eldre.beukes@aru.co.uk

Abstract

Background: Tinnitus is a symptom that can be very distressing owing to hearing sounds not related to any external sound source. Managing tinnitus is notoriously difficult, and access to evidence-based care is limited. Cognitive behavioral therapy (CBT) is a tinnitus management strategy with the most evidence of effectiveness but is rarely offered to those distressed by tinnitus. The provision of internet-based CBT for tinnitus overcomes accessibility barriers; however, it is not currently readily available in the United States.

Objective: The aim of this study is to investigate the efficacy of internet-based CBT compared with that of weekly monitoring for the management of tinnitus in reducing tinnitus distress; reducing tinnitus-related comorbidities, including tinnitus cognitions, insomnia, anxiety, and depression; and assessing the stability of the intervention effects 2 months after the intervention.

Methods: A 2-arm randomized clinical trial comparing audiologist-guided internet-based CBT (n=79) to a weekly monitoring group (n=79) with a 2-month follow-up assessed the efficacy of internet-based CBT. Eligible participants included adults seeking help for tinnitus. Recruitment was conducted on the web using an open-access website. Participants were randomized via 1:1 allocation, but blinding was not possible. The study was undertaken by English or Spanish speakers on the web. The primary outcome was a change in tinnitus distress as measured using the Tinnitus Functional Index. Secondary outcome measures included anxiety, depression, insomnia, tinnitus cognition, hearing-related difficulties, and quality of life.

Results: Internet-based CBT led to a greater reduction in tinnitus distress (mean 36.57, SD 22) compared with that in weekly monitoring (mean 46.31, SD 20.63; effect size: Cohen *d*=0.46, 95% CI 0.14-0.77) using an intention-to-treat analysis. For the secondary outcomes, there was a greater reduction in negative tinnitus cognition and insomnia. The results remained stable over the 2-month follow-up period. No important adverse events were observed. Further, 16% (10/158) of participants withdrew, with low overall compliance rates for questionnaire completion of 72.3% (107/148) at T1, 61% (91/148) at T2, and 42% (62/148) at T3.

Conclusions: This study is the first to evaluate and indicate the efficacy of audiologist-delivered internet-based CBT in reducing tinnitus distress in a US population. It was also the first study to offer internet-based CBT in Spanish to accommodate the large Hispanic population in the United States. The results have been encouraging, and further work is indicated in view of making such an intervention applicable to a wider population. Further work is required to improve compliance and attract more Spanish speakers.

Trial Registration: ClinicalTrials.gov NCT04004260; https://clinicaltrials.gov/ct2/show/NCT04004260

(J Med Internet Res 2022;24(2):e27584) doi:10.2196/27584

KEYWORDS

tinnitus; cognitive behavioral therapy; internet intervention; web-based intervention; randomized controlled trial; telehealth; teleaudiology; eHealth

Introduction

Background

Tinnitus is a condition characterized by the perception of sound in the absence of an external stimulus. It is highly prevalent, with at least 10% of Americans experiencing some form of tinnitus, of which a proportion have chronic burdensome or debilitating tinnitus [1]. Bothersome tinnitus causes various functional impairments in sleep, concentration, cognitive performance, and thought processing [2,3]. It is also associated with an increased risk of psychological difficulties, such as anxiety, depression, and reduced quality of life [4,5]. This may result in participant restrictions and activity limitations [6,7]. Owing to the negative impact of tinnitus, those distressed by their tinnitus require interventions to help them cope with the condition.

Managing tinnitus is notoriously challenging because it is often not a curable medical cause [8]. When comorbid problems such as hearing loss accompany tinnitus, hearing aids and sound therapy may reduce the severity of tinnitus [8]. Although these interventions may reduce tinnitus perception, they do not always alter negative reactions to tinnitus. Psychological interventions that change reactions to tinnitus have effectively helped reduce tinnitus distress [9]. The intervention with the strongest research evidence according to the American Academy of Otolaryngology–Head and Neck Surgery tinnitus practice guidelines [10-12] and several systematic reviews [9,13] is cognitive behavioral therapy (CBT) for tinnitus. CBT is a psychological intervention that addresses unhelpful thought patterns and emotional reactions caused by tinnitus [14]. Despite the evidence base, accessibility to CBT for tinnitus is limited owing to a dearth of health care providers with the knowledge and expertise to provide CBT to this population [15,16].

Previous Work

To overcome this barrier, an internet-based CBT for tinnitus [17] has been developed. This intervention was originally developed in Swedish [18] and was later translated into German [19] and English [20] and was provided with psychological guidance. To further increase accessibility, internet-based CBT for tinnitus has been adapted to be delivered by audiologists [21] with some training to handle the CBT elements without compromising outcomes [22-27]. The efficacy of internet-based CBT has been indicated in 9 clinical trials across mainland Europe and the United Kingdom (for a review, see the study by Beukes et al [28]). However, no clinical trials determined the effects of internet-based CBT in the United States. An evidence-based, standardized approach, such as internet-based CBT, is desirable, as tinnitus provision varies substantially across clinics and providers [16].

Study Rationale

To address this need, internet-based CBT was adapted for the US population to improve cultural and linguistic suitability [29]. It was further translated into Spanish to serve the large Spanish-speaking population in the United States, and functionality and acceptability testing was undertaken [30]. Internet-based CBT for tinnitus in the United States was evaluated within a clinical trial framework to evaluate complex interventions [31]. A small (N=28) phase I trial was undertaken in the United States [32], indicating feasibility. However, a larger randomized controlled trial (RCT) is needed to determine its efficacy in the US population. Efficacy cannot be assumed owing to many differences in health care provision in the United States and Europe. In the United States, tinnitus interventions include Tinnitus Retraining Therapy [33] and Progressive Tinnitus Management [34] with limited provision of CBT for tinnitus. It is also not known if a psychological approach will be acceptable because of the large emphasis of most tinnitus management programs on sound therapy and the fitting of devices [35]. The fee structure for health care in the United States is also very different to largely free of charge health care in the United Kingdom, as it is generally paid out of pocket, and clinicians have great difficulty receiving payment for nondiagnostic appointments such as tinnitus counseling. Before the COVID-19 pandemic, health care was also generally provided face to face. Hence, the uptake for a remote intervention is uncertain but has now become a more urgent matter. These factors may be potential barriers to or facilitators of internet-based CBT in the United States.

This RCT aims to explore the effects of internet-based CBT in the United States with the following:

1. To evaluate the efficacy of audiologist-delivered internet-based CBT in reducing tinnitus distress compared with that in weekly monitoring of tinnitus.
2. To ascertain the efficacy of internet-based CBT in reducing comorbidities associated with tinnitus.
3. To assess the stability of internet-based CBT intervention effects 2 months after the intervention.

The hypothesis was that patients with tinnitus would experience greater reduction of tinnitus distress and comorbidities after receiving internet-based CBT compared with patients receiving weekly monitoring.

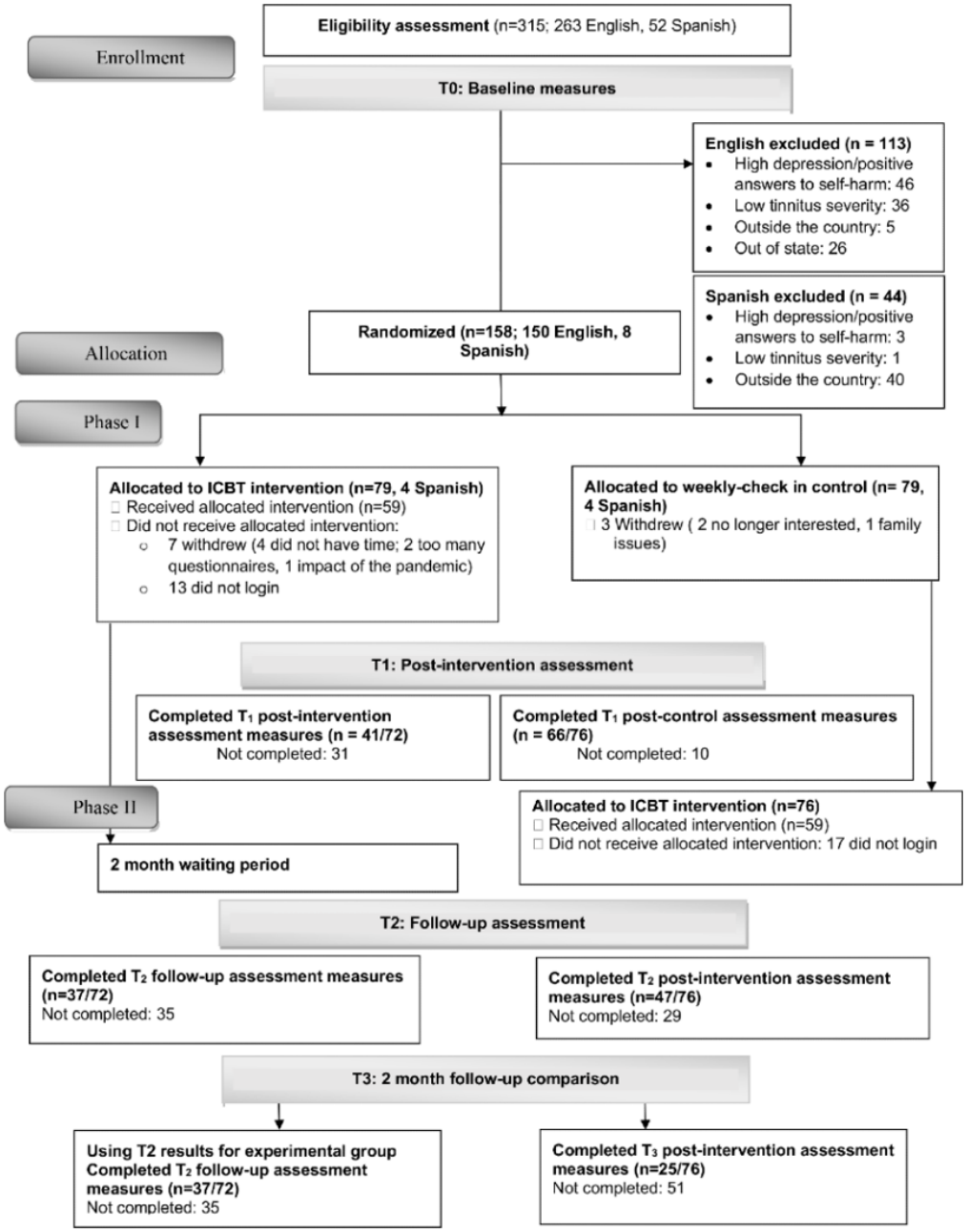
Methods

Trial Design

Overview

A prospective 2-arm delayed intervention efficacy trial with a 2-month follow-up was conducted. As an efficacy trial, an active comparator was not included. Participants were randomized with a 1:1 allocation ratio to the experimental group to receive the internet-based CBT intervention for 8 weeks or the control group whose participants were monitored weekly during the 8-week period. During the first phase, the experimental group completed the intervention. Following the experimental group completing the interventions, both groups completed the same outcome measure (T1, post intervention). During the second phase, the control group underwent the same internet-based CBT intervention, after which both groups were invited to complete the T2 outcomes, which were completed after the control group finished the intervention and 2-months postintervention for the experimental group (T2). This study design, therefore, provided the opportunity to evaluate the intervention effects in 2 independent groups at 3 time points as shown in Figure 1.

Figure 1. The CONSORT (Consolidated Standards of Reporting Trials) flow diagram. ICBT: internet-based cognitive behavioral therapy.



Ethics and Preregistration

This RCT and protocol were preregistered at ClinicalTrials.gov (trial number: NCT04004260) on July 2, 2019. Ethical approval was obtained from the Institutional Review Board at Lamar University, Beaumont, Texas, United States (IRB-FY17-209). The study was conducted and reported according to the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) guidelines [36] (Multimedia Appendix 1). An independent data-monitoring committee monitored the trial. There were no changes in the methods or trial outcomes used after the trial commenced. The participants were closely monitored for any harm. No harm or unintended effects have been reported.

Participants

The study was undertaken on the web and no clinical visits were required. The study eligibility criteria are presented in Textbox 1.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Adults, aged ≥18 years, living in Texas, United States
- The ability to read and type in English or Spanish
- Having access to a computer and the internet and able to email
- Experiencing tinnitus for a minimum period of 3 months
- A tinnitus severity score of ≥25 on the Tinnitus Functional Index indicates the need for an intervention

Exclusion criteria

- Indications of significant depression (≥15) on the Patient Health Questionnaire-9
- Indications of self-harm thoughts or intent, answer affirming question 10 of the Patient Health Questionnaire-9
- Reporting any medical or psychiatric conditions that could interfere with the treatment
- Reporting pulsatile, objective, or unilateral tinnitus, which has not been investigated medically or tinnitus still under medical investigation
- Undergoing any tinnitus therapy concurrent with participation in this study

Eligibility was determined by a 2-stage process as follows:

1. A web-based screening questionnaire, which included demographic information, health- and mental health–related questions, and standardized outcome measures are shown in Table 1.
2. A telephone interview during which the researcher rechecked eligibility and provided the opportunity for potential participants to ask any questions related to the study. The study procedures were explained, and motivational interviewing was conducted to encourage participants to commit and engage in the intervention.
3. Participants who did not meet the inclusion criteria were participants with a score of ≥15 on the Patient Health Questionnaire-9 or indicated self-harm on question 10 received a phone consultation from a clinical psychologist on the research team. This call ensured that they were under care elsewhere, or necessary resources or referrals were provided.

Table 1. Study outcome measures used before and after the intervention and at the 2-month follow-up period.

Dimension	Outcome measures	Internal consistency	Scores (range)	Levels of significance	Time frame measured
Primary outcome measure					
Tinnitus distress	Tinnitus Functional Index: 42	0.97	0-100; a reduction in scores indicates improvement	>25 indicates a mild condition (no need for an intervention), 26-50 indicates a significant condition (possible need for an intervention), and ≥50 indicates a severe condition (need for a more intense intervention)	T0[a], T1[b], T2[c], and T3[d] (control only)
Secondary outcome measures					
Generalized anxiety	Generalized Anxiety Disorder-7: 44	0.89	0-21; a reduction in scores indicates improvement	0-4 indicates minimal anxiety, 5-9 indicates mild anxiety, 10-14 indicates moderate anxiety, and 5-21 indicates severe anxiety	T0, T1, T2, and T3 (control only)
Depression	Patient Health Questionnaire-9: 45	0.83	0-27; a reduction in scores indicates improvement	5-9 indicates mild depression, 10-14 indicates moderate depression, 15-19 indicates moderately severe depression, and 18-20 indicates severe depression	T0, T1, T2, and T3 (control only)
Insomnia	Insomnia Severity Index: 46	0.74	0-28; a reduction in scores indicates improvement		T0, T1, T2, and T3 (control only)
Tinnitus cognitions	Tinnitus Cognitions Questionnaire: 47	0.91	0-104; a reduction in scores indicates improvement	Higher scores indicate a greater tendency to engage in negative cognitions in response to tinnitus	T0, T1, T2, and T3 (control only)
Health-related quality of life	EQ-5D-5L[e]: 48	0.7-0.85	0-15; a reduction in scores indicates improvement	Measures 5 dimensions: mobility, self-care, usual activities, pain, discomfort, anxiety and depression	T0, T1, T2, and T3 (control only)
Health-related quality of life	EQ-5D-5L visual analog scale: 48	0.7-0.85	0-100; higher scores indicate improved health	Visual analog scale for overall health	T0, T1, T2, and T3 (control only)
Short measure for tinnitus, hearing disability, and hyperacusis	Tinnitus and Hearing Survey: 49	0.86-0.94	Subscale for tinnitus: 0-16; subscale for hearing: 0-16; subscale for sound tolerance: 0-8	—[f]	T0, T1, T2, and T3 (control only)
Weekly monitoring					
Screening of tinnitus severity	Tinnitus Handicap Inventory–Screening [37]	0.93	0-40; a reduction in scores indicates improvement	>6 indicates tinnitus handicap	Weekly, while undertaking the 8-week intervention

Dimension	Outcome measures	Internal consistency	Scores (range)	Levels of significance	Time frame measured
Tinnitus perception	Tinnitus Qualities Questionnaire [21]	Not assessed	0-100; a reduction in scores indicates improvement	Designed to determine whether tinnitus qualities such as loudness, pitch, the number of tones heard, and so forth improves while undertaking an intervention. Higher scores indicate that more bothersome aspects of tinnitus are present.	Weekly, while undertaking the 8-week intervention

[a]T0: before intervention.

[b]T1: 8 weeks after the experimental group started the intervention (before the control group started the intervention).

[c]T2: 8 weeks after the control group completed the intervention (2 months after the experimental group completed the intervention).

[d]T3: 2 months after the control group completed the intervention.

[e]EQ-5D-5L: European Quality of Life Five Dimension.

[f]Higher scores indicate more problematic tinnitus, hearing disability, and hyperacusis.

Recruitment Strategy

In line with the US government's health promotion initiative to make health care linguistically and culturally accessible [38], all the study materials were available in both English and Spanish and were also written at or below the sixth-grade English reading level [29,39]. The participants were mostly recruited from the general public using a range of strategies including a television broadcast, promoting the study via tinnitus support groups in Texas and the American Tinnitus Association, and contracting the company TrialFacts to boost recruitment. Further recruitment strategies included the use of social media (eg, Facebook and Twitter), flyers, and posters, which were distributed to local communities and put up in clinic waiting rooms. Professionals such as audiologists and otolaryngologists in Texas were also notified about the study and provided with leaflets to distribute to suitable patients. Recruitment was conducted on the web from an open access website between February 17, 2020, and March 30, 2020. Those interested were directed to the study website, which had detailed information about the study and the study team and registered their interest in study participation. Informed consent was provided on the web, confirming the understanding of how data would be used, to be randomized, the length of the trial, the commitment expected, and being contacted for follow-up data collection. Following registration, they were invited to complete the web-based screening, demographic, and outcome questionnaire. They were informed of their right to withdraw without penalty at any stage of the process.

Sample Size, Power, and Attrition

Sample size estimation was calculated using G*Power (version 3.1.6) [40] and based on achieving a 13-point clinically meaningful change between the baseline and postintervention measurements using the primary assessment measure, the Tinnitus Functional Index (TFI). Pilot data [32] indicated 26 participants per group with a 1:1 allocation to achieve 80% power to detect a between-group mean standardized difference effect size of Cohen d=0.50 (a moderate effect size). As this was fewer than the 58 participants suggested using data from an internet-based CBT RCT from the United Kingdom [23], we selected 58 participants per group. In addition, the sample size was inflated to account for missing data, estimated to be 20% of the US phase I trial data [32]. Thus, 146 participants were recruited, with 73 in each arm (calculated as 58/0.8).

Randomization

Participants meeting the inclusion criteria were randomly assigned in a 1:1 ratio and enrolled in either the experimental or control group using a computer-generated randomization schedule by an independent research assistant in blocks of varying sizes after participants were prestratified for language (English and Spanish). Participants and investigators could not be blinded to the group allocation owing to the nature of the intervention. Participants were informed immediately after randomization when the intervention commenced by the principal investigator, but not explicitly to which group they were assigned.

Patient-Public Partnership

A patient-public partnership was established to include 2 individuals with tinnitus who had piloted the internet-based CBT intervention, 2 audiologists, and 2 researchers. The meetings were held via videoconferencing. The aim of the patient-public partnership was to guide the study processes and provide inputs to the research strategy to boost recruitment and other elements of the study to ensure high compliance and engagement

Intervention

The internet-based CBT intervention content was based on a CBT self-help program originally developed in Swedish [18] and later adapted and translated into English [20]. The intervention was subsequently transformed into an 8-week interactive e-learning version suitable for the UK population [41] and then adapted linguistically and culturally to ensure suitability for the US population [29]. These adaptions prioritized accessibility of the intervention, such as lowering the readability to below the recommended sixth-grade level, as more than half of the US adult population has low literacy skills [42]. The intervention was augmented by a mindfulness module and more videos. The internet-based CBT platform (version 1)

in the US application consisted of 22 modules with worksheets and quizzes (see [21,29] for more details). Participants required an internet connection to access the materials and email correspondence regarding the intervention.

The intervention platform was transferred from Sweden and housed in the United States (Lamar University) to comply with the required US data protection regulations. Before this feasibility trial, the acceptability and functionality of this intervention for the US population were first ensured [30].

Both groups received the same intervention, and only the timing of receiving the intervention varied. The control group received the experimental intervention 8 weeks after the experimental group commenced the program. The intervention was 8 weeks long. Participants were asked to read the modules weekly and ideally spent 10 minutes each day practicing the suggested strategies.

Audiology Guidance

Guidance was provided to support the participants while undertaking the intervention. This included monitoring progress, monitoring weekly scores, providing feedback on worksheets completed, outlining the content of new modules, and answering questions. Participants who did not engage were contacted to support their participation and to discuss possible barriers. An encrypted 2-way messaging system was used to communicate with a minimum guidance time of 10 minutes per participant. Although psychologists have traditionally provided CBT interventions, audiologists generally deliver tinnitus management [16]. Thus, audiologists provided guidance to participants in a manner consistent with previous English trials [22-24] using this intervention to ensure standardization of the intervention approach. Support to the Spanish participants was provided by a Doctor of Audiology student with clinical experience whose first language was Spanish, using a handbook that was developed by the lead English-speaking audiologist.

Outcome Measures

Primary Outcome Measure

The primary outcome measure was tinnitus severity as measured by the TFI [43]. A meaningful change was defined to occur when scores were reduced by ≥13 points [43]. The TFI has been translated into >15 languages and has been validated for several populations, including the Chinese, Dutch, Swedish, and German [44]. It was selected over other tinnitus questionnaires as it was specifically developed to measure tinnitus severity and assess responsiveness to treatment and for comparison with previous trials [22-25].

Secondary Outcome Measures

Secondary outcome measures assessed anxiety [45], depression [46], insomnia [47], tinnitus cognitions [48], general health-related quality of life [49], and a short measure of hearing-related difficulties [50] as shown in Table 1. All questionnaires were used with the required permissions, and agreements were set up for those that were not freely available to use. For Spanish speakers, validated Spanish-translated versions were used. Where these were unavailable, validated translations were undertaken [39].

Weekly Monitoring of Tinnitus During the Active Intervention Period

Throughout the program, tinnitus in participants was monitored weekly using the Tinnitus Handicap Inventory–Screening (THI-S) version. The THI-S consists of a 10-item questionnaire, and scores are comparable (r=0.90) with the full version of the Tinnitus Handicap Inventory (THI) [37]. Weekly scores were also used to detect possible adverse effects. If scores increased by more than 10 points between 2 consecutive weeks, this was noted as an adverse effect. Those indicating adverse effects were contacted to address the identified problems. Participants were also monitored using a newly developed Tinnitus Qualities Questionnaire (TQQ) [21]. The TQQ measures tinnitus qualities such as pitch, loudness, and the number of tones heard. The TQQ scores can range from 0 to 100, with higher scores indicating more problematic tinnitus.

Intervention Variables

Intervention compliance was assessed by determining the retention rates and compliance in completing the outcome questionnaires. Intervention engagement was assessed by the number of log-ins, the number of modules read, and the number of messages sent during the intervention. Adverse effects were monitored by (1) direct questioning in the outcome questionnaire regarding the presence of adverse effects, (2) adverse effects written in messages or worksheets, and (3) an increase of 10 points or more during weekly monitoring using the THI-S questionnaire.

Questionnaire Administration

Web-based questionnaires were used throughout the study for both groups. Although not all measures were validated for web-based use, the results were to be comparable, as equivalent psychometric properties have been reported [51]. All the measures were completed at baseline (T0), T1 (after the intervention for experimental group), T2 (after the intervention for the control group) and at the 2-month follow-up for the experimental group). To measure the control group 2 months after the intervention, participants from the control group completed further outcome measures at T3 (adjusted). For data analysis, the T3 results for the control group and those at T2 for the experimental group were compared to assess the invention effect at the same experimental time point for both groups (ie, 2 months after the intervention). To maximize retention, 3 electronic reminders were sent to participants who had not completed questionnaires on 3 consecutive days after the release of the questionnaire. A further reminder was sent via email and text messages. If questionnaires were not completed, participants were telephoned to encourage completion of the questionnaire. Participants were also telephoned after completing the intervention to discuss the progress they had made and share their questionnaire results.

Statistical Analysis Plan

Statistical analyses were performed using the SPSS (version 26.0). All statistical tests were 2-tailed with an α set to .05. To account for missing data from participants not completing the postintervention or follow-up intervention analysis, an imputation analysis was undertaken. As the data were missing

at random, missing data were handled through multiple imputation using the Markov Chain Monte Carlo approach owing to its ability to reduce bias even when the proportion of missing data was large [52,53]. For comparison, a complete case analysis was also performed by analyzing only the completed questionnaire data without imputing missing data. As there were substantial differences, statistical analysis using the imputed data is reported in the results section as a more accurate unbiased account of the findings.

The primary study outcome was a change in the TFI score among the groups at after the intervention (T1). A difference in scores between T1 and T2 for the experimental group was used to assess the stability of the intervention effects. Effect sizes, linear mixed effects models, and a reliable change index were used to assess the primary and secondary outcomes. Changes in baseline to postintervention scores were compared within and among the groups using the pre- and posttest effect sizes (Cohen *d*) for all primary and secondary outcomes using the observed data. Effect sizes of Cohen d=0.20 represent small effect sizes, those of Cohen d=0.50 represent medium effect sizes, and those Cohen $d \geq 0.80$ represent large effect sizes [54].

A linear mixed model, which provided unbiased results in the presence of missing data (using all available data), was applied to analyze the intervention effect over time for each outcome measure. An unstructured repeated effects and identify-random effects covariance structure provided the best model fit based on the Akaike Information Criterion. Time was treated as a repeated and fixed effect. A restricted maximum-likelihood estimation was applied. The Type III *F* test sums of squares from the linear mixed model were calculated. As a sensitivity analysis, baseline tinnitus severity was initially added as a covariate, but as it had no significant effect on the results, it was removed from the model.

Another model was run to test the differences during the course of the 8-week intervention for weekly tinnitus outcome measures. Posthoc time comparisons were carried out in the case of significant group differences to assess at which time points these differences occurred. In addition to statistical significance, clinical significance has also been reported. A 13-point difference is recommended by the original developers of the TFI [43] to indicate a meaningful change in scores. To handle study variability, the reliable change index [55] is recommended as a means of calculating clinical significance for the TFI as the primary outcome. This was calculated using the mean pretest–posttest score difference, the pretreatment SD (17.49), and a test–retest reliability coefficient of 0.78 as reported in the validation study.

Sample Characteristics

Descriptive statistics including gender, age, ethnicity, race, tinnitus duration, hearing aid use, professional consultations, ease of computer use, veteran status, education, and employment status were used to describe the sample. The means and SDs were reported for each outcome measure at each time point. Descriptive statistics, including the number of log-ins and modules read, were also used to describe the sample and intervention engagement. A chi-square test of independence was used to identify group differences in engagement and compliance rates.

Results

Participant Characteristics

A total of 158 (50.2%) adults of the 315 participants screened met the eligibility criteria and were randomly assigned to the experimental (n=79) and control groups (n=79) as shown in Figure 1. In all, 8 (5.1%; 4 in each group) of the 158 participants were Spanish speakers who completed the Spanish version of the internet-based CBT program. Of the total sample (N=158), 80 participants (50.6%) were women and 78 participants (49.4%) were men with a mean age of 57 (SD 12) years for the full sample. The groups were well matched, and there were no clinically meaningful differences as seen in Table 2. Most participants (144/158, 91.1%) indicated that they were frequent users of computers and the internet. There were no functional failures related to the intervention during the trial. This trial commenced at the end of March 2020. This timing was unfortunate, as it coincided with the peak of the COVID-19 pandemic. Some participants reported becoming ill, struggling to adjust emotionally, or finding the required lifestyle changes difficult.

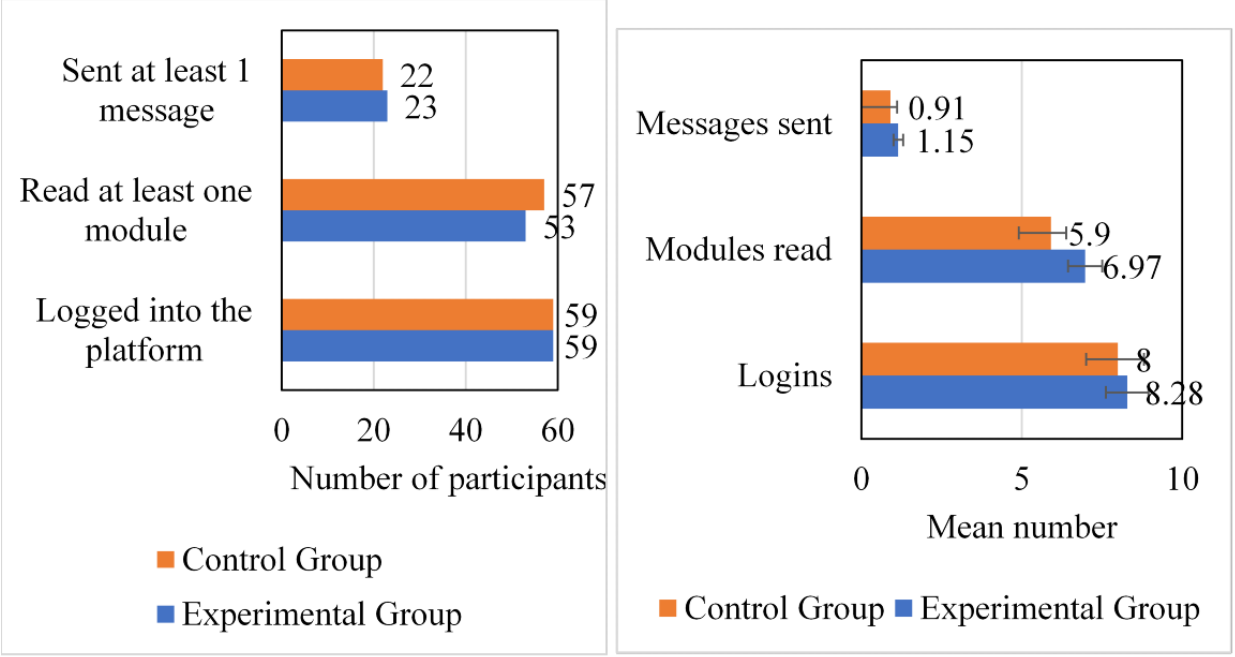
Table 2. Demographical characteristics of the participants.

Category and description	Experimental group (n=79)	Control group (n=79)	Overall (N=158)
Gender, n (%)			
Male	40 (50.6)	38 (48.1)	78 (49.3)
Female	39 (49.4)	41 (51.9)	80 (50.6)
Age (years), mean (SD; range)	56 (13; 19-76)	58 (11; 29-84)	57 (12; 19-84)
Tinnitus duration (years), mean (SD; range)	15 (16; 4 months to 70 years)	12 (12; 3 months to 58 years)	14 (14; 3 months to 70 years)
Ethnicity, n (%)			
Hispanic or Latino	9 (11.4)	11 (13.9)	20 (12.7)
Not Hispanic or Latino	70 (88.6)	68 (86.1)	138 (87.3)
Race, n (%)			
American Indian or Alaska native	0 (0)	0 (0)	0 (0)
Asian	1 (1.3)	0 (0)	1 (0.5)
Native Hawaiian or Pacific Islanders	0 (0)	0 (0)	0 (0)
Black or African American	2 (2.5)	2 (2.5)	4 (2.5)
White	74 (93.7)	74 (93.7)	148 (93.7)
More than 1 race	2 (2.5)	3 (3.5)	5 (3.1)
Highest educational level, n (%)			
High school	11 (13.9)	10 (12.7)	21 (13.2)
College or vocational training	22 (27.8)	31 (39.2)	53 (33.5)
University degree	46 (58.2)	38 (48.1)	84 (53.2)
Employment, n (%)			
Skilled or professional	55 (69.6)	41 (51.9)	96 (60.7)
Retired	22 (27.8)	30 (38)	52 (32.9)
Not working	2 (2.5)	8 (10.1)	10 (6.3)
All professionals seen, n (%)			
Primary care physician	41 (51.9)	44 (55.7)	85 (53.7)
Ear, nose, and throat physician	33 (41.8)	36 (45.6)	69 (43.7)
Audiologist	36 (45.6)	39 (49.4)	75 (47.5)
Veterans			
Veterans, n (%)	8 (10.1)	11 (13.9)	19 (12)
Duration in the military service, mean (SD; range)	8 (3; 2-10)	8 (6; 2-23)	8 (5; 2-23)
Ease of using a computer, n (%)			
Basic skills	7 (8.9)	7 (8.9)	14 (8.9)
Frequent user	72 (91.1)	72 (91.1)	144 (91.1)

Retention, Compliance, Engagement, and Adverse Effects

Overall compliance for completing the outcome measures was low, with 57% (41 participants) and 51% (37 participants) completion rates at T1 and T2, respectively, for the experimental group (Figure 1). Although compliance was greater in the control group with 87% (66 participants) and 62% (47 participants) completion at T1 and T2, there was only 33% (25 participants) completion for the control group at T3, resulting in a significant difference between the group completion rates (χ^2_3=7.98; N=411; *P*=.046) at T3 with lower completion by the control group as shown in Figure 2.

Figure 2. Intervention engagement.



Adverse effects have been reported to be low. During the intervention period, only 1 (0.6%) participant had an increase of more than 10 points in the THI-S questionnaire. On finding out more, this was related to a particularly stressful deadline for work under difficult circumstances during the COVID-19 pandemic. Only 1 (0.6%) participant reported an adverse effect on the outcome questionnaire, explaining that initially their tinnitus was more bothersome because of the focus on tinnitus at the start of the intervention. There were no serious adverse events, such as privacy breaches or major technical problems.

Intervention engagement was low but varied considerably among the participants. To identify if group allocation contributed, engagement among groups was compared, whereas each group was actively involved with the intervention. However, no significant group differences were identified (χ^2_1=0.13; N=273; *P*=.93) as shown in Figure 2. On average, 82% (59/72 participants) of the experimental group and 77% (59/76 participants) of the control group logged into the platform; 74% (53/72 participants) from the experimental group and 79% (57/72 participants) of the control group read at least one module, and 32% (23/72 participants) from the experimental group and 30% (23/76 participants) from the control group sent at least one message.

Efficacy of Internet-Based CBT in Reducing Tinnitus Distress Compared With Weekly Monitoring

The tinnitus severity among the treatment arms was not constant over time (Figure 3 and Multimedia Appendix 2). The mean difference indicated a significant difference for the internet-based CBT group with an effect size of Cohen *d*=0.46 at T1. The test of fixed effects (Table 3) indicated that the intercept, slope, and group by time interaction had significant effects on the changes in tinnitus severity. There was no estimated difference in baseline tinnitus severity among the groups (*P*=.92). The posttreatment effect was significantly lower in the control group. After the experimental group underwent treatment (T1), they had an estimated 10-point decrease in tinnitus severity (CI 3-16; $t_{156.00}$=−2.88; *P*=.004). After the control group also underwent treatment (T2), there was a nonsignificant estimated difference of 5 points (CI −2 to 11; $t_{156.00}$=−1.48; *P*=.14). This may have been because of the initial large reduction (mean 7.62, SD 18.21 points) in scores during weekly monitoring despite not having the intervention.

Figure 3. Change in tinnitus severity among groups over time. T1: only the experimental group had the intervention; T2: after the intervention for the control group and 2-month follow-up for the experimental group; T3: the comparison of the 2-month follow-up for both groups.

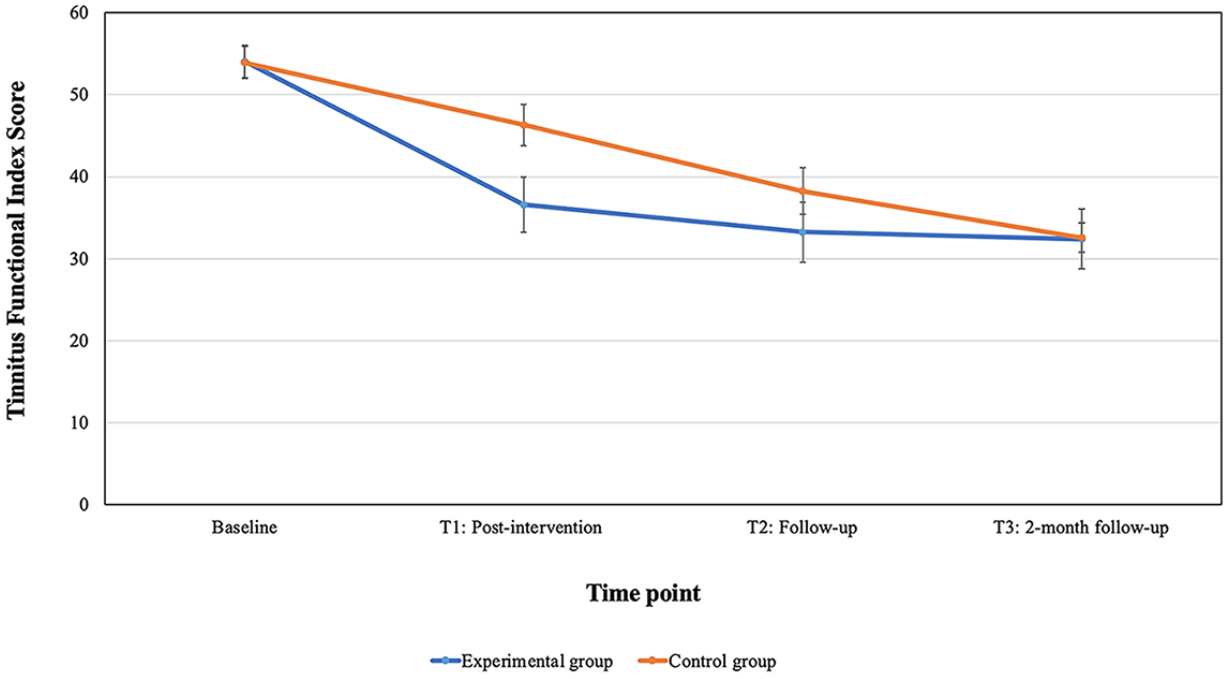


Table 3. Random intercept mixed model results using results from the imputation data.

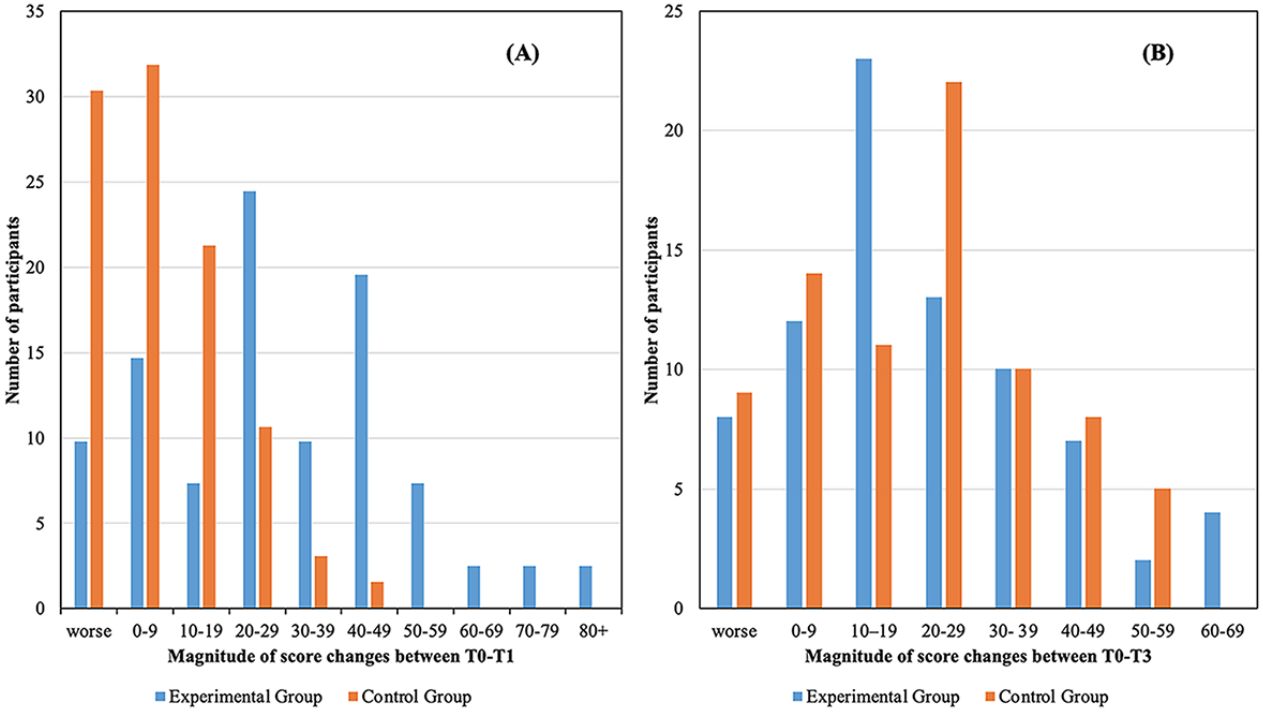
Outcome predictor	Intercept		Time		Group		Time × group	
	F test (*df*)	*P* value	*F* test (*df*)	*P* value	*F* test (*df*)	*P* value	*F* test (*df*)	*P* value
Tinnitus	1364.769 (1,156)	<.001	77.21 (3,156)	<.001	2.80 (1,156)	.10	3.64 (3,156)	.01
Anxiety	534.153 (1,156)	<.001	4.74 (3,156)	.003	0.05 (1,156)	.83	0.841 (3,156)	.47
Depression	489.593 (1,156)	<.001	12.250 (3,156)	<.001	0.05 (1,156)	.82	0.163 (3,156)	.92
Insomnia	637.397 (1,156)	<.001	42.064 (3,156)	<.001	1.81 (1,156)	.18	2.33 (3,156)	.08
EQ-5D-5L[a]	2034.549 (1,156)	.001	14.33 (3,156)	<.001	0.13 (1,156)	.72	0.19 (3,156)	.90
EQ-5D-5L visual analog scale	8100.537 (1,156)	<.001	2.02 (3,156)	.11	0.64 (1,156)	.43	1.63 (3,156)	.19
Tinnitus and Hearing Survey: tinnitus	294.231 (1,156)	<.001	38.850 (3,156)	<.001	0.750 (1,156)	.39	3.312 (3,156)	.02
Hearing disability	526.930 (1,156)	<.001	21.511 (3,156)	.001	2.24 (1,156)	.14	2.15 (3,156)	.10
Hyperacusis	247.016 (1,156)	<.001	2.51 (3,156)	.06	0.017 (1,156)	.90	0.410 (3,156)	.75
Tinnitus cognitions	1715.178 (1,156)	<.001	19.29 (3,156)	<.001	2.28 (1,156)	.13	4.15 (3,156)	.007

[a]EQ-5D-5L: European Quality of Life Five Dimension.

The model indicated an estimated baseline to 2-month follow-up mean difference of 22 points (CI 18-25) after undertaking the intervention with an estimated TFI score of 32 (baseline score was 54) at follow-up (CI 30-34).

A comparison of the margin of score reduction between T0 and T1 is shown in Figure 4. This indicates that the experimental group had a greater score reduction (between 20 and 50 points) owing to the intervention with a maximum reduction of 88 points, compared with a maximum reduction of 44 points in the control group (with the majority between 0 and 9 points) who were only monitored weekly during this period.

Figure 4. (A) Magnitude of Tinnitus Functional Index score changes between T0 and T1 after the experimental group underwent internet-based cognitive behavioral therapy. The control group was monitored weekly. (B) The magnitude of these changes between T0 and T3 at the 2-month follow-up after both groups completed the full intervention.



The clinical significance was calculated using a reliable change index. The reliable change criterion was calculated to be 22.74 in TFI score. Using this value, clinical significance was achieved by 45 (57%) participants in the experimental group and 12 (15%) participants from the control group at T1 (after the intervention was performed in the experimental group). A clinically significant change was found in 40 (51%) participants of the experimental group at T2 (2 months after the intervention) and 30 (38%) participants of the control group (after the control group completed the intervention) and 40 (48%) participants of the control group at the 2-month follow-up.

Efficacy of Internet-Based CBT in Reducing Tinnitus Comorbidities Compared With Weekly Monitoring

Results from the secondary assessment measures among the treatment arms were not constant over time, except for hyperacusis and the European Quality of Life Five Dimension visual analog scale scores for health-related quality of life, which did not have significant time effects (Table 3). After the intervention (T1), the experimental group had a significantly greater reduction in tinnitus cognition scores, indicating a medium effect (rounded off to Cohen *d*=0.50). The intercept, slope, and time by group interaction revealed significant effects on the changes in tinnitus cognition. Baseline tinnitus cognition was not significantly different among the groups (*P*=.58). After completing the intervention (T1), the experimental group had an estimated 8-point decrease in tinnitus cognition (CI 3-13; t_{156}=−3.13; *P*=.002). After both groups undertook the intervention (T2), there was no significant estimated difference among the groups (*P*=.08). For tinnitus cognitions, the model indicated an estimated baseline to 2-month follow-up mean difference of 7 points (CI 3-11) after undertaking the intervention with an estimated score of 32 at follow-up (CI 31-32).

The experimental group had a significantly greater reduction in insomnia scores after the intervention (T1) and at follow-up (T2), although there was only a small effect (Cohen *d*=0.30) at T1 and medium effect (rounded off Cohen *d*=0.50) at T2. There was no significant group by time interaction indicated by the test of fixed effects owing to later improvements by the control group. Likewise, although the experimental group had a significantly greater reduction in hearing disability at T1, this was only a small effect (Cohen *d*=0.3) without a significant group by time interaction owing to later improvements by the control group.

Confirming the results of the primary outcome, the Tinnitus and Hearing Survey secondary measure indicated that the experimental group had a significantly greater reduction in tinnitus (Cohen *d*=0.60) with no significant difference at baseline or after the control group completed the intervention. The test of fixed effects results indicated that there was no significant time by group interaction for anxiety, depression, health-related quality of life, or hyperacusis outcome measures, and no significant effect was seen.

Stability of Internet-Based CBT Intervention Effects 2 Months After the Intervention

At the 2-month follow-up, the experimental group indicated further reduction in tinnitus severity and all other outcomes. There were no significant differences in scores between T1 and T2 for the experimental group, indicating that intervention effects were maintained 2 months after the intervention.

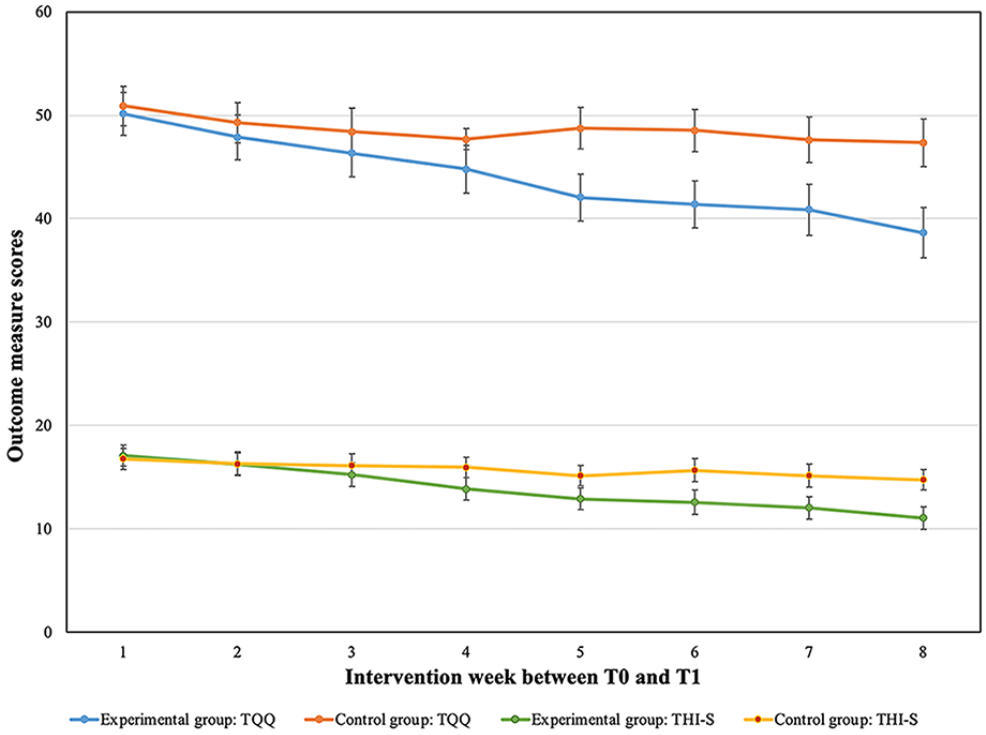
At the 2-month follow-up, most of the secondary outcome measure scores were stable. However, there was an increase in

anxiety scores for both groups, an increase in negative tinnitus cognitions for the experimental group, and a decrease in health-related quality of life for the control group, although these differences were not statistically significant.

Comparison of Weekly Tinnitus Severity During the Active Intervention Period

Differences among the intervention arms were not constant across the eight time points between T0 and T1 for both the THI-S and TQQ outcome measure scores. The experimental group had a greater weekly reduction in tinnitus distress as evidenced by the significant group by time interaction for both the THI ($F_{7,136.000}$=4.02; P=.04) and TQQ ($F_{7,136.000}$=2.55; P=.02), as well as a significant intercept and slope. Pairwise comparisons indicated significant differences among groups from week 5 for the TQQ and week 6 for the THIs, as the experimental group's (receiving internet-based CBT) tinnitus distress was rated significantly lower than that of the control group (not undergoing internet-based CBT). The maximum between-group mean difference in scores was at week 8, with the experimental group having a THI-S score of 3.5 (SE 1.1) points lower and TQQ score of 8.72 (SE 2.30) lower than that of the control group, as seen in Figure 5.

Figure 5. Weekly monitoring between the experimental and control group between T0 and T1. THI-S: Tinnitus Handicap Inventory–Screening; TQQ: Tinnitus Qualities Questionnaire.



Discussion

Overview

This study is the first to evaluate the efficacy of an audiologist-delivered internet-based CBT in reducing tinnitus distress in a US population. It was also the first study to offer internet-based CBT in Spanish to accommodate the Hispanic population in the United States. The study objectives were to evaluate the efficacy of the audiologist-delivered internet-based CBT in reducing tinnitus distress and the comorbidities associated with tinnitus compared with that of weekly monitoring of tinnitus. Furthermore, the stability of the intervention effects was assessed 2 months after the intervention.

Principal Findings

Participating in the internet-based CBT intervention led to significantly greater improvements in tinnitus distress and a medium effect size compared with that of weekly monitoring. This adds to the evidence base regarding the feasibility of audiologist-guided internet-based CBT as indicated in clinical trials in the United Kingdom using audiologist guidance [22-25]. These results are also in line with earlier internet-based CBT trials for tinnitus, indicating a pooled medium effect size (Cohen d=0.59) (for a review see the study by Beukes et al [28]) with those from Europe using psychologist guidance.

Most participants had a reduction of between 20 and 50 points in their TFI scores, although a range of outcomes was observed. Improvements were found in the control group after weekly monitoring, which has also been previously mentioned [56]. This may also be the effect of the control group knowing that they will be receiving an intervention, and this expectation may have helped them manage better before receiving the actual intervention. This initial improvement did seem to affect the postintervention results, which were lower in the control group than in the experimental group. When comparing the participants in each group on a weekly basis, it was observed that the experimental group had a greater reduction in tinnitus distress over the 8-week period. Significant differences were present from week 5 for TQQ and week 6 for THI. This was slightly later in the intervention than was found in the United Kingdom

trial when comparing differences for the THI by week 4 [23]. After the control group undertook the intervention, they made similar significant improvements as those demonstrated by the experimental group, and no significant differences were found among the groups. These results were maintained at the 2-month follow-up for both groups, although the magnitude of reduction was more variable, with the most participants from the experimental group indicating a 10- to 19-point reduction and those from the control group, a 20- to 29-point reduction in TFI scores. Further studies are required to assess whether they are maintained long-term (eg, 1 year) as has been found by previous internet-based CBT for tinnitus trials [17,25,57,58].

More participants from the experimental group had a clinically significant difference (57%; 45 participants) after their treatment compared with 38% (30 participants) of the control group following their treatment. At the 2-month follow-up, 51% (40 participants) and 48% (40 participants) from each group achieved a clinically significant change in tinnitus distress. This was lower than for the pilot study [32] owing to the reliable change criterion required being higher for this study because of a larger baseline SD. As the reliable change criterion was similar to internet-based CBT efficacy trial in the United Kingdom [23], comparable proportions of participants reached clinical significance in this study.

Secondary Results

Experiencing tinnitus is accompanied by various comorbidities that may exacerbate distress and negative emotional responses to perception. An important aspect of tinnitus intervention is the ability to address these challenging comorbidities. Undergoing internet-based CBT resulted in a significantly greater reduction in negative tinnitus cognitions (Cohen *d*=0.46) and insomnia (small effect). This finding is in line with the pooled results from previous internet-based CBT studies, also indicating a small effect (Cohen *d*=0.42). This was the first internet-based CBT for tinnitus study that included the tinnitus cognition questionnaire. The finding that this CBT intervention is able to reduce negative thought patterns associated with tinnitus is a positive finding and continued use of the tinnitus cognition questionnaire is important as recommended by Handscomb et al [59]. From the pooled results of internet-based CBT studies, there was no effect on quality of life [28], similar to this study, but there was a greater reduction in anxiety and depression, which was not found in this study. The reason may be related to the exclusion of individuals with severe mental health conditions, possibly reducing the opportunity to observe an intervention effect owing to the low baseline scores. Broader inclusion criteria are necessary to ensure internet-based CBT is provided to all affected individuals, including those with mental health conditions, as they seem to benefit from this as shown in previous studies [17,19,21-25,57,58]. The pilot study also indicated an effect for hearing disability and hyperacusis [32], which was not observed in this study. The content addressed in the modules providing hearing tactic strategies and advice on reducing sound sensitivity were in the optional modules, which were not read by many participants, and this may have contributed to these results. Interestingly, changes were noted in the TQQ, suggesting that the internet-based CBT may result in a change in tinnitus perception (eg, tinnitus pitch, loudness, and number of sounds heard) in addition to a reduction in tinnitus distress. However, these observations should be replicated in future studies, in addition to possible biomarkers.

Comparison With Previous Work

The study participants' characteristics were similar to those found in previous internet-based CBT trials; however, the mean age was slightly higher at 57 years than at a mean age of 51 years [28]. Despite extensive recruitment strategies and campaigns, following suggestions to support Hispanic and Latino research participants [60], only 8 participants selected to participate in the intervention in Spanish. Further efforts will be required to build trust within Hispanic communities before recruiting for subsequent trials [61].

Although the study administration mimicked that that would be provided in a routine application, the overall completion rate of the posttreatment questionnaires was low across time points and groups, compared with that of previous internet-based CBT for tinnitus studies [28]. Although only 10 participants (6%) withdrew, many enrolled participants never logged into the intervention website. Moreover, not all the modules were read, and very few messages were sent, indicating low intervention engagement. Numerous factors contribute to this finding. One may be the timing of this study during the COVID-19 pandemic. Participants explained that they were on their computers all day attending meetings on Zoom as they had to stay at home. This may be an additional computer work, making this intervention difficult, because of them wanting to take a break from their computers. Some participants mentioned having contracted the COVID-19 virus, and even after recovering, they remained fatigued, making intervention engagement difficult. Others found the lifestyle changes of working from home and juggling childcare difficult, and some were affected emotionally. However, the COVID-19 pandemic is unlikely to be the only reason for poor engagement. Differences in compliance among groups were also observed. After receiving the intervention, compliance was lower in the experimental group and lower in the control group at the 2-month follow-up. This may have reflected the trial design that the control group had an additional assessment time point and having already completed assessments after the control and intervention, decided not to complete another assessment at the 2-month follow-up. During the pilot study [32], lower engagement than noted in earlier studies was observed. Cultural differences may not be accounted for. In contrast to the United Kingdom and many parts of mainland Europe, many people pay for health care via third-party reimbursements in the United States. The internet-based CBT intervention was offered free of charge. It may be that people undervalued this treatment as a free treatment that may be perceived as less effective than one requiring payment. In addition, this study recruited participants only from Texas, United States, thus representing a very small population of the United States and may not represent the wider US population. Subsequent trials should be performed with a wider US population. Process evaluation may be helpful in identifying the factors contributing to the retention and engagement rates and identifying what may improve theses [26]. Continued public involvement in planning and implementing subsequent research

trials will be vital to gain insights into the factors important to participants [62,63].

Limitations

This study represents participants living in Texas, United States, who do not present with severe mental health conditions, often associated with tinnitus. This may not present the typical tinnitus population, and thus, the findings cannot be generalized to other populations. Despite recruitment efforts, only 8 Spanish participants were recruited; similarly, the participant groups also contained low numbers of ethnic and racial minorities when compared with the general population from this region. The study was furthermore conducted during the peak of the COVID-19 pandemic, a time when day-to-day living was disrupted for most people. The results may also have been different if a large proportion of participants were engaged and completed the outcome assessments.

Conclusions

Compounding the potentially debilitating nature of severe tinnitus, accessible, evidence-based interventions are still lacking. There is an urgent need to improve the availability of such interventions. Furthermore, the COVID-19 pandemic has highlighted the need for evidence-based eHealth approaches to overcome the limited in-person contact and support available for individuals with tinnitus [63]. These results further support the role of audiologists in guiding such forms of tinnitus management. The results have been encouraging, and further work is indicated in view of making such an intervention applicable to a wider population. Future work should consider enrolling heterogeneous tinnitus populations to examine who are more suitable (or not) for the internet-based CBT program. In addition, a stepped approach (eg, a brief intervention offered to all participants and a comprehensive intervention for more suitable patients following the brief intervention) should be examined to improve compliance and engagement.

Acknowledgments

We would like to thank webmasters George Vlaescu and Shrinivas Varadaraj for their help in installing the intervention platform and for providing technical assistance. We extend our thanks to all the participants who participated in this study; our public-patient advisory group; and the organizations, including the American Tinnitus Association, that helped with recruitment. We also thank Professor Stacia DeSantis for the statistical assistance provided.

This work was funded by the National Institute on Deafness and Other Communication Disorders of the National Institute of Health under award number R21DC017214. The funding body was not involved in the design of the study, data collection, data analysis, interpretation of data, or writing of the manuscript.

Authors' Contributions

VM obtained the funding and licenses. GA provided the software. All authors contributed to the research design. Data collection, analysis, and interpretation were performed by EB. EB drafted the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-eHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1093 KB - jmir_v24i2e27584_app1.pdf]

Multimedia Appendix 2
Outcome measures at each time point.
[DOCX File , 18 KB - jmir_v24i2e27584_app2.docx]

References

1. Spankovich C, Gonzalez VB, Su D, Bishop CE. Self reported hearing difficulty, tinnitus, and normal audiometric thresholds, the National Health and Nutrition Examination Survey 1999-2002. Hear Res 2018 Feb;358:30-36. [doi: 10.1016/j.heares.2017.12.001] [Medline: 29254853]
2. Clarke NA, Henshaw H, Akeroyd MA, Adams B, Hoare DJ. Associations between subjective tinnitus and cognitive performance: systematic review and meta-analyses. Trends Hear 2020;24:1-23 [FREE Full text] [doi: 10.1177/2331216520918416] [Medline: 32436477]
3. Watts EJ, Fackrell K, Smith S, Sheldrake J, Haider H, Hoare DJ. Why is tinnitus a problem? A qualitative analysis of problems reported by tinnitus patients. Trends Hear 2018;22:1-10 [FREE Full text] [doi: 10.1177/2331216518812250] [Medline: 30482102]

4. Salazar JW, Meisel K, Smith ER, Quiggle A, McCoy DB, Amans MR. Depression in patients with tinnitus: a systematic review. Otolaryngol Head Neck Surg 2019 Jul;161(1):28-35 [FREE Full text] [doi: 10.1177/0194599819835178] [Medline: 30909841]
5. Trevis KJ, McLachlan NM, Wilson SJ. A systematic review and meta-analysis of psychological functioning in chronic tinnitus. Clin Psychol Rev 2018 Mar;60:62-86. [doi: 10.1016/j.cpr.2017.12.006] [Medline: 29366511]
6. Beukes EW, Manchaiah V, Andersson G, Allen PM, Terlizzi PM, Baguley DM. Situationally influenced tinnitus coping strategies: a mixed methods approach. Disabil Rehabil 2018 Dec;40(24):2884-2894. [doi: 10.1080/09638288.2017.1362708] [Medline: 28793790]
7. Manchaiah V, Beukes EW, Granberg S, Durisala N, Baguley DM, Allen PM, et al. Problems and life effects experienced by tinnitus research study volunteers: an exploratory study using the ICF classification. J Am Acad Audiol 2018;29(10):936-947. [doi: 10.3766/jaaa.17094] [Medline: 30479266]
8. Zenner H, Delb W, Kröner-Herwig B, Jäger B, Peroz I, Hesse G, et al. A multidisciplinary systematic review of the treatment for chronic idiopathic tinnitus. Eur Arch Otorhinolaryngol 2017 May;274(5):2079-2091. [doi: 10.1007/s00405-016-4401-y] [Medline: 27995315]
9. Landry EC, Sandoval XC, Simeone CN, Tidball G, Lea J, Westerberg BD. Systematic review and network meta-analysis of cognitive and/or behavioral therapies (CBT) for tinnitus. Otol Neurotol 2020 Feb;41(2):153-166. [doi: 10.1097/MAO.0000000000002472] [Medline: 31743297]
10. Cima RFF, Mazurek B, Haider H, Kikidis D, Lapira A, Noreña A, et al. A multidisciplinary European guideline for tinnitus: diagnostics, assessment, and treatment. HNO 2019 Mar;67(Suppl 1):10-42. [doi: 10.1007/s00106-019-0633-7] [Medline: 30847513]
11. Fuller TE, Haider HF, Kikidis D, Lapira A, Mazurek B, Norena A, et al. Different teams, same conclusions? A systematic review of existing clinical guidelines for the assessment and treatment of tinnitus in adults. Front Psychol 2017;8:206 [FREE Full text] [doi: 10.3389/fpsyg.2017.00206] [Medline: 28275357]
12. Tunkel DE, Bauer CA, Sun GH, Rosenfeld RM, Chandrasekhar SS, Cunningham ER, et al. Clinical practice guideline: tinnitus. Otolaryngol Head Neck Surg 2014 Oct;151(2 Suppl):1-40. [doi: 10.1177/0194599814545325] [Medline: 25273878]
13. Fuller T, Cima R, Langguth B, Mazurek B, Vlaeyen JW, Hoare DJ. Cognitive behavioural therapy for tinnitus. Cochrane Database Syst Rev 2020 Jan 08;1:CD012614 [FREE Full text] [doi: 10.1002/14651858.CD012614.pub2] [Medline: 31912887]
14. Andersson G. Psychological aspects of tinnitus and the application of cognitive-behavioral therapy. Clin Psychol Rev 2002 Sep;22(7):977-990. [doi: 10.1016/s0272-7358(01)00124-6] [Medline: 12238249]
15. Bhatt JM, Lin HW, Bhattacharyya N. Prevalence, severity, exposures, and treatment patterns of tinnitus in the United States. JAMA Otolaryngol Head Neck Surg 2016 Oct 01;142(10):959-965 [FREE Full text] [doi: 10.1001/jamaoto.2016.1700] [Medline: 27441392]
16. Henry JA, Piskosz M, Norena A, Fournier P. Audiologists and tinnitus. Am J Audiol 2019 Dec 16;28(4):1059-1064. [doi: 10.1044/2019_AJA-19-0070] [Medline: 31689367]
17. Andersson G, Strömgren T, Ström L, Lyttkens L. Randomized controlled trial of internet-based cognitive behavior therapy for distress associated with tinnitus. Psychosom Med 2002;64(5):810-816. [doi: 10.1097/01.psy.0000031577.42041.f8] [Medline: 12271112]
18. Andersson G, Kaldo V. Internet-based cognitive behavioral therapy for tinnitus. J Clin Psychol 2004 Feb;60(2):171-178. [doi: 10.1002/jclp.10243] [Medline: 14724924]
19. Jasper K, Weise C, Conrad I, Andersson G, Hiller W, Kleinstäuber M. Internet-based guided self-help versus group cognitive behavioral therapy for chronic tinnitus: a randomized controlled trial. Psychother Psychosom 2014;83(4):234-246. [doi: 10.1159/000360705] [Medline: 24970708]
20. Abbott JM, Kaldo V, Klein B, Austin D, Hamilton C, Piterman L, et al. A cluster randomised trial of an internet-based intervention program for tinnitus distress in an industrial setting. Cogn Behav Ther 2009;38(3):162-173. [doi: 10.1080/16506070902763174] [Medline: 19675959]
21. Beukes E, Andersson G, Manchaiah V, Kaldo V. Cognitive Behavioral Therapy for Tinnitus. San Diego, CA: Plural Publishing; 2021.
22. Beukes EW, Allen PM, Manchaiah V, Baguley DM, Andersson G. Internet-based intervention for tinnitus: outcome of a single-group open trial. J Am Acad Audiol 2017 Apr;28(4):340-351. [doi: 10.3766/jaaa.16055] [Medline: 28418328]
23. Beukes EW, Baguley DM, Allen PM, Manchaiah V, Andersson G. Audiologist-guided internet-based cognitive behavior therapy for adults with tinnitus in the United Kingdom: a randomized controlled trial. Ear Hear 2018;39(3):423-433. [doi: 10.1097/AUD.0000000000000505] [Medline: 29095725]
24. Beukes EW, Andersson G, Allen PM, Manchaiah V, Baguley DM. Effectiveness of guided internet-based cognitive behavioral therapy vs face-to-face clinical care for treatment of tinnitus: a randomized clinical trial. JAMA Otolaryngol Head Neck Surg 2018 Dec 01;144(12):1126-1133 [FREE Full text] [doi: 10.1001/jamaoto.2018.2238] [Medline: 30286238]
25. Beukes EW, Allen PM, Baguley DM, Manchaiah V, Andersson G. Long-term efficacy of audiologist-guided internet-based cognitive behavior therapy for tinnitus. Am J Audiol 2018 Nov 19;27(3S):431-447 [FREE Full text] [doi: 10.1044/2018_AJA-IMIA3-18-0004] [Medline: 30452747]

26. Beukes EW, Manchaiah V, Baguley DM, Allen PM, Andersson G. Process evaluation of internet-based cognitive behavioural therapy for adults with tinnitus in the context of a randomised control trial. Int J Audiol 2018 Feb;57(2):98-109. [doi: 10.1080/14992027.2017.1384858] [Medline: 28990807]
27. Beukes EW, Manchaiah V, Davies AS, Allen PM, Baguley DM, Andersson G. Participants' experiences of an internet-based cognitive behavioural therapy intervention for tinnitus. Int J Audiol 2018 Dec;57(12):947-954. [doi: 10.1080/14992027.2018.1514538] [Medline: 30295113]
28. Beukes EW, Manchaiah V, Allen PM, Baguley DM, Andersson G. Internet-based interventions for adults with hearing loss, tinnitus, and vestibular disorders: a systematic review and meta-analysis. Trends Hear 2019;23:2331216519851749 [FREE Full text] [doi: 10.1177/2331216519851749] [Medline: 31328660]
29. Beukes EW, Fagelson M, Aronson EP, Munoz MF, Andersson G, Manchaiah V. Readability following cultural and linguistic adaptations of an internet-based intervention for tinnitus for use in the United States. Am J Audiol 2020 Jun 08;29(2):97-109 [FREE Full text] [doi: 10.1044/2019_AJA-19-00014] [Medline: 32105512]
30. Manchaiah V, Vlaescu G, Varadaraj S, Aronson EP, Fagelson MA, Munoz MF, et al. Features, functionality, and acceptability of internet-based cognitive behavioral therapy for tinnitus in the United States. Am J Audiol 2020 Sep 03;29(3):476-490 [FREE Full text] [doi: 10.1044/2020_AJA-20-00002] [Medline: 32880499]
31. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. Br Med J 2000 Sep 16;321(7262):694-696 [FREE Full text] [doi: 10.1136/bmj.321.7262.694] [Medline: 10987780]
32. Beukes E, Andersson G, Fagelson M, Manchaiah V. Audiologist-supported internet-based cognitive behavioral therapy for tinnitus in the United States: a pilot trial. J Med Internet Res 2021;In Press. [doi: 10.2196/27584]
33. Jastreboff PJ, Jastreboff MM. Tinnitus Retraining Therapy (TRT) as a method for treatment of tinnitus and hyperacusis patients. J Am Acad Audiol 2000 Mar;11(3):162-177. [Medline: 10755812]
34. Henry J, Zaugg T, Myers P, Kendall C. Progressive Tinnitus Management - Clinical Handbook for Audiologists. San Diego, CA: Plural Pub Inc; 2010:50-97.
35. Tyler RS, Perreau A, Powers T, Watts A, Owen R, Ji H, et al. Tinnitus sound therapy trial shows effectiveness for those with tinnitus. J Am Acad Audiol 2020 Jan;31(1):6-16. [doi: 10.3766/jaaa.18027] [Medline: 31210635]
36. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
37. Newman CW, Sandridge SA, Bolek L. Development and psychometric adequacy of the screening version of the tinnitus handicap inventory. Otol Neurotol 2008 Apr;29(3):276-281. [doi: 10.1097/MAO.0b013e31816569c4] [Medline: 18277308]
38. Healthy People 2010. Centers for Disease Control and Prevention. 2015. URL: https://www.cdc.gov/nchs/healthy_people/hp2010.htm [accessed 2022-01-09]
39. Manchaiah V, Muñoz MF, Hatfield E, Fagelson MA, Aronson EP, Andersson G, et al. Translation and adaptation of three English tinnitus patient-reported outcome measures to Spanish. Int J Audiol 2020 Jul;59(7):513-518 [FREE Full text] [doi: 10.1080/14992027.2020.1717006] [Medline: 31994949]
40. Faul F, Erdfelder E, Buchner A, Lang A. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 2009 Nov;41(4):1149-1160. [doi: 10.3758/BRM.41.4.1149] [Medline: 19897823]
41. Beukes EW, Vlaescu G, Manchaiah V, Baguley DM, Allen PM, Kaldo V, et al. Development and technical functionality of an Internet-based intervention for tinnitus in the UK. Internet Interv 2016 Nov;6:6-15 [FREE Full text] [doi: 10.1016/j.invent.2016.08.002] [Medline: 30135809]
42. National assessment of adult literacy. National Center for Education Statistics. 2003. URL: https://nces.ed.gov/naal/ [accessed 2022-01-09]
43. Meikle MB, Henry JA, Griest SE, Stewart BJ, Abrams HB, McArdle R, et al. The tinnitus functional index: development of a new clinical measure for chronic, intrusive tinnitus. Ear Hear 2012;33(2):153-176. [doi: 10.1097/AUD.0b013e31822f67c0] [Medline: 22156949]
44. Henry JA, Griest S, Thielman E, McMillan G, Kaelin C, Carlson KF. Tinnitus Functional Index: development, validation, outcomes research, and clinical application. Hear Res 2015 Jun 12:58-64. [doi: 10.1016/j.heares.2015.06.004] [Medline: 26074306]
45. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
46. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. J Am Med Assoc 1999 Nov 10;282(18):1737-1744. [doi: 10.1001/jama.282.18.1737] [Medline: 10568646]
47. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med 2001 Jul;2(4):297-307. [doi: 10.1016/s1389-9457(00)00065-4] [Medline: 11438246]
48. Wilson PH, Henry JL. Tinnitus Cognitions Questionnaire: development and psychometric properties of a measure of dysfunctional cognitions associated with tinnitus. Int Tinnitus J 1998;4(1):23-30. [Medline: 10753381]

49. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
50. Henry JA, Griest S, Zaugg TL, Thielman E, Kaelin C, Galvez G, et al. Tinnitus and hearing survey: a screening tool to differentiate bothersome tinnitus from hearing difficulties. Am J Audiol 2015 Mar;24(1):66-77 [FREE Full text] [doi: 10.1044/2014_AJA-14-0042] [Medline: 25551458]
51. Thorén ES, Andersson G, Lunner T. The use of research questionnaires with hearing impaired adults: online vs. paper-and-pencil administration. BMC Ear Nose Throat Disord 2012;12:12 [FREE Full text] [doi: 10.1186/1472-6815-12-12] [Medline: 23107440]
52. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. BMC Med Res Methodol 2017 Dec 06;17(1):162 [FREE Full text] [doi: 10.1186/s12874-017-0442-1] [Medline: 29207961]
53. Madley-Dowd P, Hughes R, Tilling K, Heron J. The proportion of missing data should not be used to guide decisions on multiple imputation. J Clin Epidemiol 2019 Jun;110:63-73 [FREE Full text] [doi: 10.1016/j.jclinepi.2019.02.016] [Medline: 30878639]
54. Cohen J. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [doi: 10.1037//0033-2909.112.1.155] [Medline: 19565683]
55. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J Consult Clin Psychol 1991 Feb;59(1):12-19. [doi: 10.1037//0022-006x.59.1.12] [Medline: 2002127]
56. Phillips JS, McFerran DJ, Hall DA, Hoare DJ. The natural history of subjective tinnitus in adults: a systematic review and meta-analysis of no-intervention periods in controlled trials. Laryngoscope 2018 Jan;128(1):217-227. [doi: 10.1002/lary.26607] [Medline: 28425615]
57. Kaldo V, Levin S, Widarsson J, Buhrman M, Larsen H, Andersson G. Internet versus group cognitive-behavioral treatment of distress associated with tinnitus: a randomized controlled trial. Behav Ther 2008 Dec;39(4):348-359. [doi: 10.1016/j.beth.2007.10.003] [Medline: 19027431]
58. Weise C, Kleinstäuber M, Andersson G. Internet-delivered cognitive-behavior therapy for tinnitus: a randomized controlled trial. Psychosom Med 2016 May;78(4):501-510. [doi: 10.1097/PSY.0000000000000310] [Medline: 26867083]
59. Handscomb LE, Hall DA, Shorter GW, Hoare DJ. Positive and negative thinking in tinnitus: factor structure of the Tinnitus Cognitions Questionnaire. Ear Hear 2017;38(1):126-132 [FREE Full text] [doi: 10.1097/AUD.0000000000000365] [Medline: 27560491]
60. Sha M, McAvinchey G, Quiroz R, Moncada J. Successful techniques to recruit Hispanic and Latino research participants. Surv Pract 2017 Jun 01;10(3):1-9. [doi: 10.29115/sp-2017-0014]
61. Levkoff S, Sanchez H. Lessons learned about minority recruitment and retention from the Centers on Minority Aging and Health Promotion. Gerontologist 2003 Feb;43(1):18-26. [doi: 10.1093/geront/43.1.18] [Medline: 12604742]
62. Staniszewska S, Brett J, Simera I, Seers K, Mockford C, Goodlad S, et al. GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research. Br Med J 2017 Aug 02;358:j3453 [FREE Full text] [doi: 10.1136/bmj.j3453] [Medline: 28768629]
63. Beukes EW, Baguley DM, Jacquemin L, Lourenco MP, Allen PM, Onozuka J, et al. Changes in tinnitus experiences during the COVID-19 pandemic. Front Public Health 2020;8:592878 [FREE Full text] [doi: 10.3389/fpubh.2020.592878] [Medline: 33251179]

Abbreviations

CBT: cognitive behavioral therapy
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
RCT: randomized controlled trial
TFI: Tinnitus Functional Index
THI: Tinnitus Handicap Inventory
THI-S: Tinnitus Handicap Inventory–Screening
TQQ: Tinnitus Qualities Questionnaire

Edited by R Kukafka, G Eysenbach; submitted 30.01.21; peer-reviewed by M Masalski, L Rusu; comments to author 19.03.21; revised version received 29.03.21; accepted 05.10.21; published 14.02.22.

Please cite as:
W Beukes E, Andersson G, Fagelson M, Manchaiah V
Internet-Based Audiologist-Guided Cognitive Behavioral Therapy for Tinnitus: Randomized Controlled Trial
J Med Internet Res 2022;24(2):e27584
URL: https://www.jmir.org/2022/2/e27584
doi:10.2196/27584
PMID:35156936

Original Paper

Positive Coping as a Mediator of Mobile Health Intervention Effects on Quality of Life Among People Living With HIV: Secondary Analysis of the Randomized Controlled Trial Run4Love

Yu Zeng[1,2], MB; Yan Guo[1,3], PhD; Rainbow Tin Hung Ho[4,5], PhD; Mengting Zhu[6], MS; Chengbo Zeng[7], MS; Aliza Monroe-Wise[8], MD, MSc; Yiran Li[1], MS; Jiaying Qiao[9], MS; Hanxi Zhang[10], MS; Weiping Cai[11], MD; Linghua Li[11], MD; Cong Liu[11], MSN

[1]Department of Medical Statistic, School of Public Health, Sun Yat-sen University, Guangzhou, China

[2]Longgang Center for Disease Control and Prevention in Shenzhen, Shenzhen, China

[3]Department of Population and Quantitative Health Sciences, University of Massachusetts Chan Medical School, Worcester, MA, United States

[4]Department of Social Work & Social Administration, The University of Hong Kong, Hong Kong, China

[5]Centre on Behavioral Health, The University of Hong Kong, Hong Kong, China

[6]The Jockey Club School of Public Health and Primary Care, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China

[7]Department of Global Health and Social Medicine, Harvard Medical School, Boston, MA, United States

[8]Department of Global Health, University of Washington, Seattle, WA, United States

[9]Shanghai Center for Disease Control and Prevention, Shanghai, China

[10]National Center of AIDS/Sexually Transmitted Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China

[11]Department of Infectious Diseases, Guangzhou Number Eight People's Hospital, Guangdong, China

Corresponding Author:
Yan Guo, PhD
Department of Medical Statistic
School of Public Health
Sun Yat-sen University
Guangzhou
China
Phone: 86 020 87334202
Email: Yan.Guo1@umassmed.edu

Abstract

Background: The effectiveness of psychosocial interventions on quality of life (QOL) among people living with HIV has been validated, including mobile health (mHealth) interventions. However, it is unclear which components of such interventions account for these effects.

Objective: This study aims to examine positive coping as a potential mediator of the effects of an mHealth intervention on QOL among people living with HIV.

Methods: For this secondary analysis, we used data from an mHealth-based randomized controlled trial, Run4Love, which was conducted to improve QOL and mental health outcomes of people living with HIV. A total of 300 participants were randomly assigned to the intervention group to receive the adapted cognitive-behavioral stress management courses and regular physical activity promotion or the waitlist control group in a 1:1 ratio. Our analysis focused on positive coping and QOL, which were repeatedly measured at baseline and at 3-, 6-, and 9-month follow-ups. Latent growth curve models were constructed to explore the mediating role of positive coping in the effects of the mHealth intervention on QOL.

Results: Positive coping served as a mediator in the effect of the mHealth intervention on QOL for up to 9 months. The mHealth intervention had a significant and positive indirect effect on the slope of QOL via the slope of positive coping (b=2.592×1.620=4.198, 95% CI 1.189-7.207, P=.006). The direct effect of the intervention was not significant (b=0.552, 95% CI −2.154 to 3.258, P=.69) when controlling for the mediator.

Conclusions: The longitudinal findings suggest that positive coping could be a crucial mediator of the mHealth intervention in enhancing QOL among people living with HIV. These findings underscore the importance of improving positive coping skills in mHealth interventions to improve QOL among people living with HIV.

(J Med Internet Res 2022;24(2):e25948) doi:10.2196/25948

KEYWORDS

mediation effect; mobile health; quality of life; positive coping; HIV; randomized controlled trial

Introduction

Background

Substantial improvements to and increased coverage in antiretroviral therapy have resulted in extended life expectancy of people living with HIV, leading to over 1.25 million HIV seropositive survivors in China [1,2]. With improved survival, quality of life (QOL) during treatment has become increasingly important for people living with HIV [3]. Owing to HIV disease progression and treatment side effects, the QOL of people living with HIV continues to be lower than that of people living with other chronic diseases, as well as the general population [4,5]. Poor QOL is not only related to suboptimal adherence to antiretroviral therapy, leading to inferior therapeutic effects [6], but is also associated with elevated HIV risk behavior, leading to increased risk of HIV transmission in the community [7]. Therefore, effective interventions to improve QOL in people living with HIV are warranted, especially those that can reach a large number of the targeted population, such as mobile health (mHealth) programs.

The literature suggests that psychosocial interventions such as cognitive-behavioral stress management (CBSM) programs can improve QOL in a variety of populations with chronic diseases, including people living with HIV [8-12]. Evidence also shows that mHealth-based psychosocial interventions produce similar benefits in comparison with face-to-face interventions, and they have additional advantages because of the easy accessibility and cost-effectiveness [13,14]. Nevertheless, the potential mechanism by which these mHealth interventions improve QOL among people living with HIV remains unclear. There might be various factors that affect the outcomes of the intervention. Mediation constitutes a mechanism through which the independent variable affects the dependent variable of interest [15]. Exploration of mediators will help understand the effective components of an intervention, which is crucial for future optimization of the program design [16].

A considerable amount of research has focused on exploring factors associated with QOL, with positive coping being one factor that has received substantial attention [17-21]. Positive coping is defined as individuals' active cognitive response (eg, trying to see the positive side) and adaptive behavioral response (eg, talking with family or friends) when managing stressful situations [22]. A substantial body of cross-sectional studies have shown that higher positive coping is related to higher QOL [18-20]. Some randomized controlled trials (RCTs) have indicated that positive coping might be a key determinant of the effectiveness of psychosocial interventions in improving QOL among people living with HIV [23,24]. For example, Hansen et al [24] found that HIV-positive participants reported significant postintervention improvement in both QOL and positive coping in the cognitive-behavioral support group, as well as in the control group providing mental health sessions and psychiatric services upon individual request. Using mixed models, the study also found that positive coping was directly associated with improved QOL in multiple time points across the intervention [24]. Similarly, another study revealed that a CBSM intervention was effective in improving positive coping in HIV-seropositive men who have sex with men, and positive coping might be an important predictor of QOL improvement [23].

Previous studies exploring the mediators of the effects of interventions on improving QOL have mostly been conducted in face-to-face settings [23,25,26]. It is unclear whether the factor mediating the effects of an intervention on QOL in face-to-face settings remains effective in mHealth settings. Furthermore, existing studies on mediation analysis mostly use pre- and postintervention data, whereas interventions in repeated measure design are lacking [11,23,25]. As 2–time point data contain limited information on individual changes, longitudinal data tracking over multiple time points (≥3) may allow further investigation on the changes of both the mediator and the outcome, as well as the trajectory of the time-varying relationships between them over time [27].

Objective

To bridge the gaps in the existing literature, we used longitudinal data from the Run4Love study to examine the mediating role of positive coping on QOL in an mHealth intervention. The Run4Love study was a WeChat (Tencent)-based RCT to examine the intervention effects of an adapted CBSM course with physical activity promotion compared with usual care in people living with HIV. In this study, we hypothesized that the mHealth intervention would enhance the use of positive coping strategies in people living with HIV, which in turn was related to the significant improvement in QOL over time.

Methods

Research Setting and Ethical Considerations

The study used data from an mHealth-based RCT, Run4Love (ChiCTR-IPR-17012606) [28]. The Run4Love intervention was designed to improve mental health outcomes of people living with HIV, including, but not limited to, the reduction in depressive symptoms and improvement in QOL [28,29]. It was conducted from September 2017 to October 2018 in Guangzhou, the third-largest city of China [30]. A total of 300 participants were randomly assigned to either the intervention group to receive a 3-month WeChat-based mHealth program or the waitlist control group to receive usual care in a 1:1 ratio. The study design, recruitment procedures, and process of randomization were detailed in the study protocol, which was approved by the Institutional Review Board of Sun Yat-sen University [28].

Participants Enrollment

The trained research staff recruited participants from the infectious disease outpatient department of the only hospital

designated for HIV care and treatment in Guangzhou. Patients who showed interest in the study were invited to participate in a consultation session to receive further information about the program and complete a screening questionnaire. Patients were eligible to participate if they met all of the following inclusion criteria: (1) aged ≥18 years, (2) HIV seropositive, (3) having elevated depressive symptoms (Center for Epidemiologic Studies-Depression score ≥16), and (4) using WeChat, the most popular instant messaging mobile program in China, with more than 1 billion active users worldwide [31]. Patients were excluded if they met any of the following exclusion criteria: (1) currently on psychotherapy or taking psychiatric drugs, (2) unable to complete the questionnaires, (3) unable to read or listen to the intervention materials, and (4) unable to engage in physical activities for medical reasons.

Participants who met the aforementioned eligibility criteria and provided written informed consent were enrolled in the study. Under the guidance of the research staff, participants were asked to complete electronic questionnaires in the outpatient department at baseline and 3-, 6-, and 9-month follow-ups. Participants would receive 50 RMB (ie, approximately US $8) or gifts of equivalent value (eg, a yoga mat) as an incentive for completing each assessment.

The Run4Love Intervention

Participants in the intervention group received a 3-month WeChat-based intervention, which consisted of the adapted CBSM course and physical activity promotion [28]. CBSM is a therapeutic approach to teaching people living with HIV to manage their stress by focusing on how individuals' thoughts affect their emotions and behaviors and to influencing participants' irrational thoughts and changing their thought patterns and behaviors [32]. The adapted CBSM course mainly included coping and stress reduction skills such as meditation, muscle relaxation, and abdominal respiration training. The physical activity promotion program mainly included guidance on exercise, benefits of exercise, and a healthy diet. The content of these 2 intervention components was provided through 65 intervention materials in multiple formats, including short articles, motivational posters, and audio clips. These materials were delivered to participants on a near-daily basis (5-6 times a week) via an enhanced WeChat platform, the Run4Love platform. On average, each of the short articles had around 1300 words, which required about 5 minutes to read; each poster with motivational messages required about half a minute to read; each audio recording required 5 to 10 minutes to listen to [33].

Participants in the control group received a brochure on nutrition and healthy living style. They would be offered the Run4Love program as soon as the study was completed (ie, 9 months after their enrollment).

Measurement

QOL Assessment

QOL was measured using the 31-item World Health Organization Quality of Life HIV short version (WHOQOL-HIV BREF) at baseline and 3, 6, and 9 months. The WHOQOL-HIV BREF has been widely used among people living with HIV with proven reliability and validity [34,35]. The 31-item scale measures QOL across six domains (ie, physical, psychological, level of independence, social relationships, environment, and spirituality) over the last 2 weeks. Each item was rated on a 5-point Likert scale ranging from 1 (not at all) to 5 (extremely), altogether providing a 4-20 score in each domain and a 24-120 score for the whole scale after the transformation, with higher scores indicating better QOL [34]. An example of these items assessing QOL was "How much do you enjoy life?" Good reliability of the WHOQOL-HIV BREF was shown in the study, and Cronbach α ranged from .727 to .838 across the 4 assessment points.

Positive Coping

Positive coping was measured using the 12-item subscale of the Simplified Ways of Coping Questionnaire with good reliability and validity in the Chinese populations [36-38]. The subscale consists of 12 items such as "Try to look on the bright side of things" and "Find out several different methods to solve the problem" to assess the frequency of positive attitudes and skills of coping that a participant adopted in daily life. Each item was rated on a 4-point Likert scale ranging from 0 (not use at all) to 3 (use frequently), providing a 0-36 score, with higher scores indicating higher levels of positive coping. Cronbach α ranged from .848 to .925 across the 4 assessment points in this study.

Demographic Characteristics

Demographic characteristics included age, gender, marital status, educational level, sexual orientation, employment status, family monthly income, and duration of HIV infection (years).

Statistics Analysis

The intention-to-treat principle was applied to all analyses in this study. Descriptive analyses of demographic characteristics, QOL, and positive coping were performed. Differences in outcome measures and baseline characteristics between the intervention and control groups were evaluated using the 2-tailed *t* test or Wilcoxon rank-sum test for continuous variables and the chi-square test for categorical variables. Latent growth curve models (LGCMs) were constructed to examine the mediating role of positive coping in the effects of the mHealth intervention on QOL among people living with HIV. All statistical hypothesis tests were 2-sided, and a *P* value <.05 was considered statistically significant. Descriptive analyses and group comparisons of the key variables were conducted using R (version 3.5), and LGCMs were constructed using Mplus (version 7; Muthén & Muthén). Missing data in all LGCMs were managed using the robust maximum likelihood estimation procedure.

The analyses of the mediation effect were conducted in 3 steps [39]. First, 2 unconditional LGCMs were constructed separately to estimate the growth trajectories of the outcome measure (QOL) and potential mediator variable (positive coping). The latent variables of intercept and slope (ie, the initial status and rate of change) were estimated from the repeated measures at 4 time points. For the intercept factor, the loadings across the 4 assessment points were fixed to 1. For the slope factor, the first loading was fixed to 0, the second to 1, and the other loadings were freely estimated [40]. Instead of assuming a linear

trajectory by setting the loadings onto the slope factor to 0, 1, 2, and 3, it was more reasonable to explore the potentially nonlinear trajectories by setting the third and fourth loadings free when the assumption of linear LGCM was unconfirmed [40].

Second, 2 conditional LGCMs were specified to separately examine the impact of the intervention on QOL and positive coping. The conditional models were extensions of the unconditional models to incorporate the variable of intervention group assignment (ie, intervention group=1 and control group=0) as a covariate.

Third, a parallel-process LGCM was constructed to evaluate whether the intervention was effective in improving QOL via the mediator variable of positive coping. The parallel-process model was a combination of the aforementioned 2 conditional LGCMs, which simultaneously estimated the trajectories of QOL and positive coping, and incorporated intervention group assignment as a covariate. The mediation effect was tested based on bias-corrected 95% bootstrapped CIs with a resampling of 5000 [41,42].

Model fit was evaluated using chi-square test statistics and other indexes, including the Tucker-Lewis index (TLI), the comparative fit index (CFI), the root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR). An LGCM with adequate model fit should meet the following criteria: TLI>0.90, CFI>0.90, RMSEA<0.08, and SRMR<0.08 [43-45].

Results

Baseline Characteristics

Descriptive statistics for baseline characteristics are shown in Table 1. The baseline characteristics of the participants were balanced between the intervention and control groups, except for the fact that a slightly higher proportion of heterosexual participants were allocated to the control group. The median age of the participants was 27.5 years, and the median duration of HIV infection was 1.7 years. The majority were male (277/300, 92.3%), well-educated (182/300, 60.7% with at least some college education), unmarried (262/300, 87.3%), and employed (251/300, 83.7%) and had a moderate level of income (176/300, 58.7% with family monthly income <7000 yuan [US $1100]). Up to 81.7% (245/300) of the participants were homosexual, bisexual, or uncertain of their sexuality.

Table 1. Baseline Characteristics of the participants in the Run4Love randomized controlled trial (N=300).

Variables	Total (N=300)	Intervention (n=150)	Control (n=150)	*P* value
Age (years), median (IQR)	27.5 (24.5-31.3)	27.4 (24.3-31.1)	27.8 (24.6-32.2)	.29
Gender, n (%)				.13
Male	277 (92.3)	142 (94.7)	135 (90)	
Female	23 (7.7)	8 (5.3)	15 (10)	
Educational level, n (%)				.10
>High school	182 (60.7)	98 (65.3)	84 (65)	
≤high school	118 (39.3)	52 (34.7)	66 (35)	
Marital status, n (%)				.73
Single or divorced or widowed	262 (87.3)	132 (88)	130 (86.7)	
Married	38 (12.7)	18 (12)	20 (85.3)	
Employment status, n (%)				.29
Employed	251 (83.7)	123 (82)	128 (85.3)	
Unemployed	49 (16.3)	27 (18)	22 (14.7)	
Family monthly income (yuan), n (%)				.20
≥7000 (US $1100)	124 (41.3)	68 (45.3)	56 (37.3)	
<7000 (US $1100)	176 (58.7)	82 (55.7)	94 (62.7)	
Sexual orientation, n (%)				.03
Homosexual or bisexual or uncertain	245 (71.7)	130 (86.7)	115 (76.7)	
Heterosexual	55 (18.3)	20 (13.3)	35 (23.3)	
Duration of HIV infection (years), median (IQR)	1.7 (0.6-3.7)	1.7 (0.6-4.0)	1.8 (0.6-3.9)	.62
Positive coping, mean (SD)	18.4 (5.8)	18.4 (5.5)	18.3 (6.2)	.92
Quality of life, mean (SD)	77.0 (9.2)	77.4 (9)	76.6 (9.4)	.44

Changes in Repeated Measures

Repeated measures of the outcome variable (ie, QOL) and potential mediator (ie, positive coping) at the 4 assessment points are presented in Table 2. The sample mean and 95% CIs of QOL and positive coping over 9 months for both the intervention and control groups are presented in Figure 1. Participants in the intervention group had significantly higher levels of QOL and positive coping than those in the control group at the 3-, 6-, and 9-month follow-ups, indicating significant effects of the Run4Love mHealth intervention on improving QOL and positive coping over time. Specifically, in comparison with those in the control group, participants in the intervention group also had significantly higher mean scores in all 6 domains of QOL at the 3-, 6-, and 9-month follow-ups (Figure 2; Table 3).

Table 2. Repeated measures of quality of life and positive coping of the participants in the Run4Love randomized controlled trial.

Variables	Quality of life, mean (SD)			Positive coping, mean (SD)		
	Intervention[a]	Control[b]	*P* value	Intervention[a]	Control[b]	*P* value
Baseline	77.43 (9.03)	76.59 (9.42)	.44	18.39 (5.46)	18.32 (6.15)	.92
3 months	82.54 (12.03)	76.63 (11.08)	<.001	20.79 (7.33)	17.70 (5.88)	<.001
6 months	83.51 (12.88)	76.32 (12.96)	<.001	21.03 (7.48)	17.38 (6.59)	<.001
9 months	83.48 (13.17)	76.54 (13.34)	<.001	20.95 (7.75)	18.31 (6.41)	.003

[a]For variables in the intervention group, the sample sizes were 150, 139, 132, and 133 at baseline and 3-, 6-, and 9-month follow-ups, respectively.

[b]For variables in the control group, the sample sizes were 150, 135, 133, and 127 at baseline and 3-, 6-, and 9-month follow-ups, respectively.

Figure 1. Repeated measures of quality of life (QOL) and positive coping of the participants in the intervention and control groups over time. Error bars represent 95% CIs.

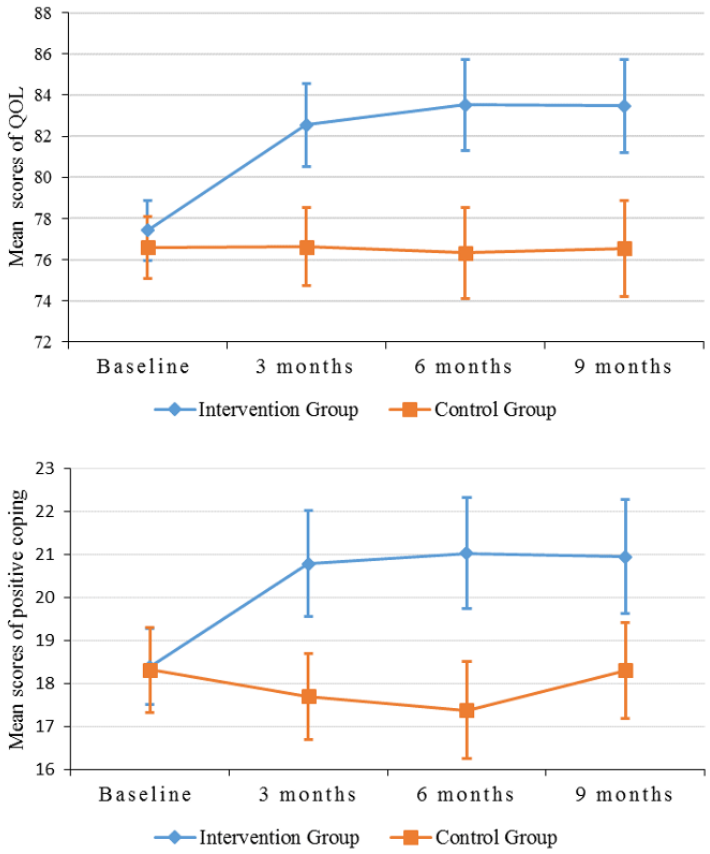


Figure 2. Repeated measures of the 6 domains in quality of life (QOL) of the participants in the intervention and control groups over time. Group 1 and Group 0 represent the intervention and control groups, respectively.

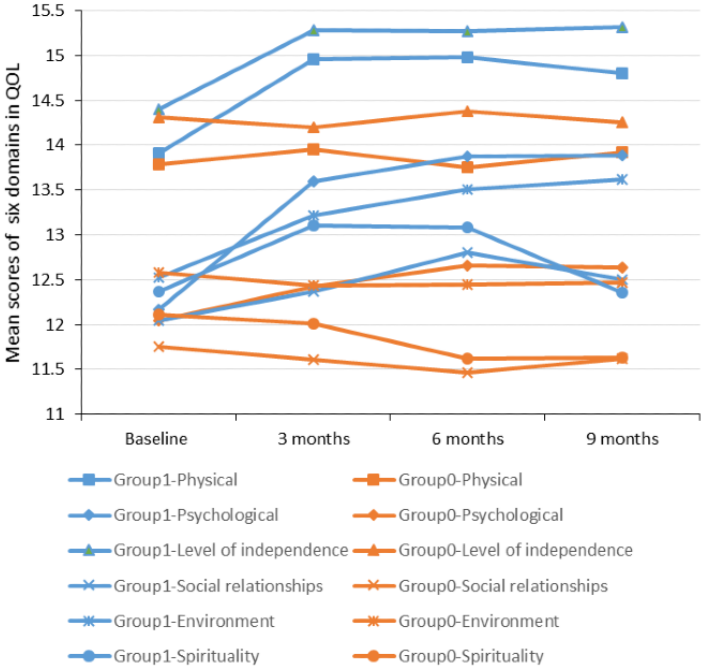


Table 3. Repeated measures of the 6 domains in quality of life of the participants in the Run4Love randomized controlled trial.

Domains in quality of life	Intervention[a]	Control[b]	*P* value
Physical			
Baseline	13.91 (2.05)	13.79 (2.39)	.66
3 months	14.96 (2.45)	13.95 (2.60)	.001
6 months	14.98 (2.69)	13.75 (2.77)	<.001
9 months	14.80 (2.84)	13.92 (3.00)	.02
Psychological			
Baseline	12.17 (2.17)	12.04 (2.08)	.60
3 months	13.60 (2.50)	12.42 (2.24)	<.001
6 months	13.88 (2.54)	12.66 (2.73)	<.001
9 months	13.89 (2.83)	12.64 (2.60)	<.001
Level of independence			
Baseline	14.40 (1.81)	14.31 (2.10)	.70
3 months	15.28 (2.04)	14.20 (2.27)	<.001
6 months	15.27 (2.23)	14.38 (2.26)	.002
9 months	15.32 (2.29)	14.26 (2.38)	<.001
Social relationships			
Baseline	12.05 (2.13)	11.75 (2.10)	.22
3 months	12.37 (2.70)	11.61 (2.22)	.01
6 months	12.80 (2.54)	11.46 (2.62)	<.001
9 months	12.50 (2.53)	11.62 (2.42)	.005
Environment			
Baseline	12.52 (1.93)	12.58 (2.08)	.80
3 months	13.22 (2.44)	12.44 (2.15)	.005
6 months	13.51 (2.45)	12.45 (2.74)	.001
9 months	13.62 (2.35)	12.47 (2.56)	<.001
Spirituality			
Baseline	12.37 (3.00)	12.11 (3.22)	.47
3 months	13.11 (3.16)	12.01 (3.29)	.005
6 months	13.08 (3.37)	11.62 (3.36)	<.001
9 months	13.36 (3.22)	11.63 (3.49)	<.001

[a]For variables in the intervention group, the sample sizes were 150, 139, 132, and 133 at baseline and 3-, 6-, and 9-month follow-ups, respectively.

[b]For variables in the control group, the sample sizes were 150, 135, 133, and 127 at baseline and 3-, 6-, and 9-month follow-ups, respectively.

The dropout rates were 8.7% (26/300; 11/150, 7.3% in the intervention group; 15/150, 10.0% in the control group), 11.7% (35/300; 18/150, 12.0% in the intervention group; 17/150, 11.3% in the control group), and 13.3% (40/300; 17/150, 11.3% in the intervention group; 23/150, 15.3% in the control group) at 3-, 6-, and 9-month follow-ups, respectively. The average completion rate of the 3-month Run4Love program among people in the intervention group was 50.8% (33/65) [33].

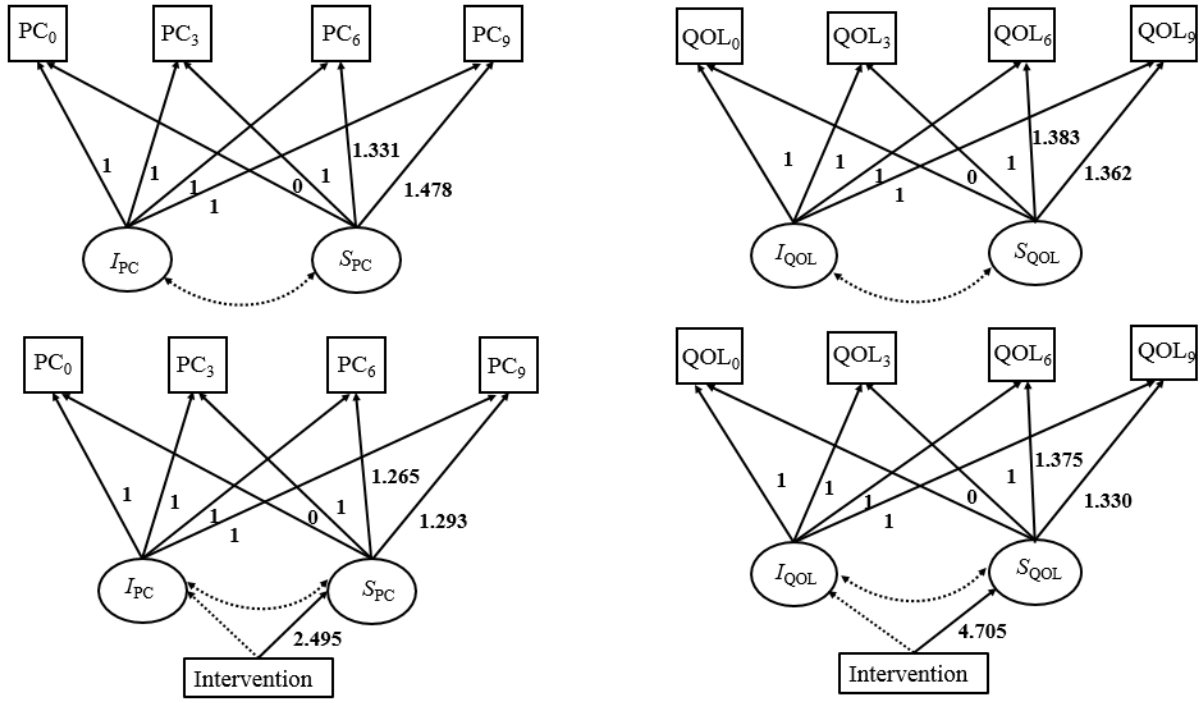
Results of LGCMs

The results of the unconditional LGCMs indicated that both QOL and positive coping improved across the course of the study, and the largest improvement occurred at the 3-month follow-up. The path diagrams of these 2 unconditional LGCMs are presented in Figure 3. The unconditional LGCM for QOL had a good model fit: χ^2_3=2.7 (*P*=.26), CFI=0.999, TLI=0.996, SRMR=0.044, and RMSEA=0.034 (Figure 3). The mean intercept was 77.032 (*P*<.001, SE 0.527), indicating a mean QOL score of 77.032 at baseline. The mean slope was 2.241 (*P*<.001, SE 0.503), and the factor loadings on the slope were 0, 1, 1.383, and 1.362 at baseline and 3, 6, and 9 months, respectively. These results reflected a nonlinear pattern of improvement in QOL: a rapid increase in QOL scores (2.241 points) occurred at 3 months, and then the magnitude of improvement flattened out. The variances of the intercept and

the slope of QOL were 92.509 (*P*=.001) and 54.013 (*P*=.04), respectively, which indicated significant individual differences in the initial levels and rates of change in QOL over time. The covariance between the intercept and the slope of QOL was −16.029 (*P*=.48), implying that the rate of change in QOL was not associated with its initial levels. The unconditional LGCM for positive coping also had a good model fit: χ^2_3=1.6 (*P*=.66), CFI=1.000, TLI=1.000, SRMR=0.025, and RMSEA<0.001 (Figure 3). The mean positive coping score was 18.373 (*P*<.001, SE=0.340) at baseline, and the mean slope was 0.699 (*P*=.04, SE=0.336). Similarly, the results of the slope factor also reflected a nonlinear pattern of improvement in positive coping: a rapid increase in positive coping scores (0.699 points) occurred at 3 months, and then the magnitude of improvement flattened out. The variances of the intercept and the slope in positive coping were 21.571 (*P*=.01) and 10.194 (*P*=.21), respectively, indicating significant individual differences in the initial levels of positive coping but not in its rates of change. The rate of change in positive coping was also unrelated to its initial levels, with a covariance between the intercept and the slope of −2.664 (*P*=.70).

The results of the conditional LGCMs indicated that the mHealth intervention had significantly positive effects on both QOL and positive coping across the course of the study. The path diagrams of these 2 conditional LGCMs are also presented in Figure 3. The conditional model for QOL fitted the data well: χ^2_5=3.3 (*P*=.50), CFI=1.000, TLI=1.000, SRMR=0.041, and RMSEA<0.001 (Figure 3). The intervention had a significant impact on the slope factor (*b*=4.705, SE=0.911, *P*<.001), indicating an improvement in QOL in the intervention group compared with the control group over time. Similarly, the conditional model for positive coping fitted the data well: χ^2_5=5.7 (*P*=.34), CFI=0.997, TLI=0.995, SRMR=0.030, and RMSEA=0.022 (Figure 3). There was a higher rate of change in positive coping in the intervention group (*b*=2.495, SE=0.826, *P*=.003), indicating a larger improvement in positive coping in the intervention group.

Figure 3. Path diagrams of the unconditional latent growth curve models (LGCMs) for quality of life (QOL) and positive coping and the conditional LGCMs with intervention groups as a covariate. Observed variables are denoted by boxes. Latent variables are denoted by ovals. Unidirectional arrows indicate the effects of 1 variable on the other. Bidirectional arrows indicate the correlations. The nonsignificant paths are shown as dotted lines. Intervention is either the Run4Love intervention group or the waitlist control group. I: intercept; PC: positive coping; S: slope; 0: baseline; 3: 3-month follow-up; 6: 6-month follow-up; 9: 9-month follow-up.



The results of the parallel process LGCM indicated that the mHealth intervention was effective in improving QOL via the mediation effect of positive coping. The path diagram of the parallel process LGCM is presented in Figure 4, and the estimates of the main coefficients are shown in Table 4. The results showed a good model fit for the parallel process LGCM: χ^2_{24}=64.1 (*P*<.001), CFI=0.971, TLI=0.957, SRMR=0.065, and RMSEA=0.075 (Figure 4). The intervention had a significantly positive impact on the slope of positive coping (*b*=2.592, *P*<.001), which in turn had a significantly positive impact on the slope of QOL (*b*=1.620, SE=0.321, *P*<.001). The indirect effect of the intervention on the slope of QOL via the slope of positive coping was significant (*b*=2.592×1.620=4.198, *P*=.006). These results indicated a mediation effect of positive coping on patients' QOL, where the exposure to the mHealth intervention significantly improved participants' positive coping over time, which in turn led to a positive change in QOL. The intervention no longer had direct effects on the slope of QOL (*b*=0.552, *P*=.69), indicating that the effects of the intervention on QOL might be explained by the mediation effect of positive coping. There were no significant differences as a function of intervention in the initial status of either repeated measures,

indicating no group differences in the starting point for either positive coping or QOL. The covariance between the 2 intercepts was 21.571 (*P*=.01), indicating that initial levels of positive coping were positively related to initial status with respect to QOL. In other words, participants who reported higher levels of positive coping at baseline tended to report higher levels of QOL at the same time. However, neither the covariance between the intercept of positive coping and the slope of QOL (2.941, *P*=.26) nor the covariance between the intercept of QOL and the slope of positive coping (−3.900, *P*=.16) was statistically significant. These results indicated that the initial level of positive coping was not significantly associated with the rate of change in QOL, nor was the initial level of QOL and the rate of change in positive coping.

Figure 4. Path diagram of a parallel process latent growth curve model for quality of life (QOL) and positive coping with intervention groups as a covariate. Observed variables are denoted by boxes. Latent variables are denoted by ovals. Unidirectional arrows indicate the effects of 1 variable on the other. Bidirectional arrows indicate the correlations. The nonsignificant paths are shown as dotted lines. Intervention is either the Run4Love intervention group or the waitlist control group. I: intercept; PC: positive coping; S: slope; 0: baseline; 3: 3-month follow-up; 6: 6-month follow-up; 9: 9-month follow-up.

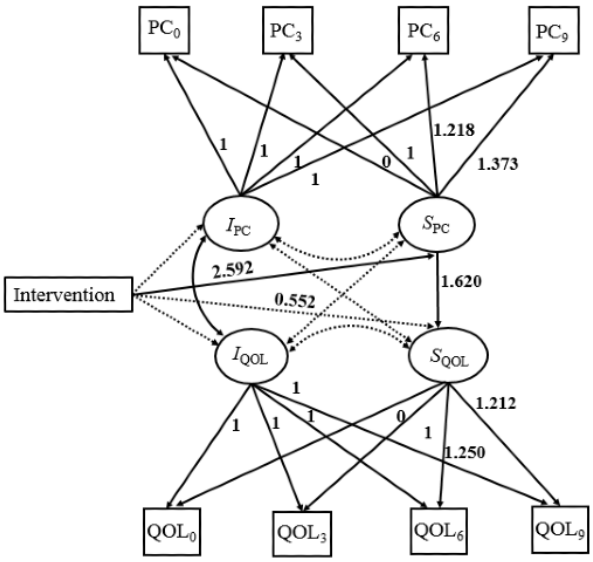


Table 4. Estimates of the coefficients in the parallel process latent growth curve model for quality of life and positive coping (n=300).

Coefficient	Estimate	95% CI	Standardized estimate	SE	*P* value
Intervention→QOL[a]	0.552	−2.154 to 3.258	0.036	1.381	.69
Intervention→positive coping	2.592	1.124 to 4.060	0.398	0.749	.001
Positive coping→QOL	1.620	0.997 to 2.243	0.692	0.318	<.001
Total effect	4.750	2.766 to 6.734	0.311	1.012	<.001
Direct effect					
Intervention→QOL	0.552	−2.154 to 3.258	0.036	1.381	.69
Indirect effect					
Intervention→positive coping→QOL	4.198	1.189 to 7.207	0.275	1.535	.006

[a]QOL: quality of life.

Discussion

Principal Findings

To the best of our knowledge, our secondary analysis of data from the Run4Love study is among the first efforts to examine the mediating role of positive coping in patients' QOL in an mHealth-based RCT among people living with HIV. The results of LGCMs demonstrated that the Run4Love trial significantly improved positive coping among people living with HIV over 9 months, and the enhancement of positive coping led to significant improvement in QOL across the study. There was full mediation between positive coping and QOL in the Run4Love trial. Thus, our study revealed one of the potential mechanisms of improvement in QOL in mHealth-based CBSM interventions. To better design and implement effective interventions for people living with HIV, more studies are needed to investigate and identify the processes and/or mechanisms by which interventions lead to significant improvements in health outcomes, especially in emerging mHealth interventions [27].

Although several cross-sectional studies and 1 cohort study have examined the mediating role of positive coping on QOL [19-21], such designs are not sufficient for a robust conclusion because their findings are based on passive observational data.

An experimental design in which the proper intervention is used to improve an individual's positive coping and QOL would greatly reduce confounding effects and enhance the strength of causal inferences. However, few RCTs have explored this relationship, especially in mHealth interventions. Only 1 face-to-face RCT showed that improvement in positive coping significantly mediated the effects of psychosocial intervention on QOL using pre- and postintervention assessments [25]. A comprehensive literature review suggested that mediation analysis could only be properly examined in well-designed RCTs with repeated measurements and sufficient power [27]. Thus, our study is well suited to investigate such a relationship and has confirmed the mediating role of positive coping on QOL, which may expand the literature on mechanisms of mHealth-based interventions.

In addition to the RCT design, the time point repeated measures and corresponding methodology of LGCM used in this study allow more conclusive findings and more abundant information for the mediation analysis or other research. Unlike cross-sectional data or pre- and postintervention data adopted in the previous literature, LGCMs that use longitudinal data allow the trajectory estimation of changes in outcome measures over time [39,40]. For example, the results of the LGCMs in our study suggested that both positive coping and QOL had increasing and nonlinear trajectories with a slower rate of increase over time, and there were statistically significant individual differences in the initial levels and rates of change of positive coping and QOL over time. In addition, the parallel process LGCM allowed for the examination of intercept-intercept, intercept-slope, and slope-slope relationships of the trajectories, which might extend our understanding of the relationships between the factors investigated [46]. In our study, we found a significant slope-slope relationship between positive coping and QOL, which indicated a mediation effect of positive coping on QOL in the mHealth intervention. Therefore, the LGCMs could not only estimate the growth curve of each outcome measure over time but also simultaneously examine the mediation relationship between 2 outcomes (eg, positive coping and QOL) in an intervention by controlling for their growth trajectories [39].

Given the critical role of positive coping in improving QOL, it is important to develop positive coping skills in mHealth interventions to improve participants' QOL. Previous psychosocial interventions for people living with HIV found that training in active cognitive and adaptive behavioral coping strategies was effective in improving positive coping [23,47,48]. In the Run4Love trial, we adapted the evidence-based CBSM intervention with important components of training in various types of coping skills, such as active cognitive coping (eg, mindfulness and problem- or emotion-oriented coping) and adaptive behavioral coping (eg, regular exercise) [28]. In addition, we adapted the CBSM courses originally designed for people living with HIV in the United States by removing some parts of the courses that were not suitable in the Chinese context, such as religion-related materials. Therefore, we had 9 sessions out of the original 10 sessions of the CBSM courses. Furthermore, as Chinese people living with HIV visit the hospital for antiretroviral drugs and regular clinical follow-ups every 3 months, we matched the patients' hospital visits and adapted the length of the intervention to 3 months by adding 3 review sessions and having weekly delivery of the total 12 sessions of the CBSM courses [28]. In addition, our CBSM intervention delivered in the mHealth platform allowed and encouraged participants' repeat visits to the training items, which might result in enhanced and sustained intervention effects over 9 months by repeatedly accessing the CBSM courses and practicing the coping skills, thus producing improved and sustained effects on patient outcomes such as QOL. To our knowledge, there is limited evidence about the long-term effects of mHealth interventions; few studies have reported intervention effects for over 6 months [49,50]. Few studies have been conducted on the potential mechanisms of the persistent intervention effects in mHealth interventions. The investigation in this study may provide important information and evidence on intervention mechanisms for long-term sustained effects in mHealth interventions.

Limitations

This study had several limitations. First, the self-reported data on positive coping and QOL in our study might have resulted in recall and social desirability biases. More objective measures, such as biomarkers, could be incorporated in future studies. Second, although participants were recruited from a large hospital for HIV treatment in Guangzhou with over 10,000 patients who were HIV seropositive, the sample was mostly from an urban setting and was predominantly male, particularly young men who have sex with men. Therefore, the generalizability of our findings should be treated with caution. Third, as all the 6 dimensions of QOL were improved in the intervention group at 3, 6, and 9 months, and the improvement was not limited to the mental health dimension, it is possible that the improvement in QOL was also related to other factors besides positive coping, such as reduced depressive symptoms, stress, and/or HIV-related stigma and increased social support [51-53]. Although it is beyond the scope of this study to incorporate other factors, one of our previous studies found mediating roles of perceived stress and depressive symptoms for suicide reduction [53], whereas another study found positive coping and HIV-related stigma as intervention mediators on depressive symptoms [52]. Therefore, other potential mediators should be further explored to clarify the mechanisms of intervention effects and processes of how mHealth interventions improve health outcomes for better designing and implementing effective mHealth interventions.

Conclusions

In conclusion, this study found a full mediation effect of positive coping on QOL among people living with HIV in an mHealth intervention using 4 time point repeated measures and LGCMs. Future research and policies aimed at QOL improvement among people living with HIV should be designed with specific features that enhance the use of positive coping strategies.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (grant 71573290) and China Medical Board open competition funding (grant 17-271). The funding sources had no role in the study design, data collection, statistical analyses, or writing of the manuscript.

Authors' Contributions

YG and YZ had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. They contributed to the conceptualization and design of the study. YZ, MZ, CZ, YL, JQ, and HZ were responsible for the acquisition, analysis, or interpretation of data. YZ drafted the manuscript. YG, RH, AMW, and YZ conducted critical revision of the manuscript for important intellectual content. LL, WC, and CL provided administrative, technical, or material support. YG supervised the study.

Conflicts of Interest

None declared.

References

1. Nakagawa F, May M, Phillips A. Life expectancy living with HIV: recent estimates and future implications. Curr Opin Infect Dis 2013 Feb;26(1):17-25. [doi: 10.1097/QCO.0b013e32835ba6b1] [Medline: 23221765]
2. HIV/AIDS in China. Wikipedia. URL: https://en.wikipedia.org/wiki/HIV/AIDS_in_China [accessed 2020-08-23]
3. Cho H, Iribarren S, Schnall R. Technology-mediated interventions and quality of life for persons living with HIV/AIDS. A systematic review. Appl Clin Inform 2017 Dec 12;8(2):348-368. [doi: 10.4338/ACI-2016-10-R-0175] [Medline: 28401246]
4. Engelhard EA, Smit C, van Dijk PR, Kuijper TM, Wermeling PR, Weel AE, et al. Health-related quality of life of people with HIV: an assessment of patient related factors and comparison with other chronic diseases. AIDS 2018 Jan 02;32(1):103-112. [doi: 10.1097/QAD.0000000000001672] [Medline: 29112062]
5. Miners A, Phillips A, Kreif N, Rodger A, Speakman A, Fisher M, ASTRA (Antiretrovirals, Sexual Transmission and Attitudes) Study. Health-related quality-of-life of people with HIV in the era of combination antiretroviral treatment: a cross-sectional comparison with the general population. Lancet HIV 2014 Oct;1(1):32-40 [FREE Full text] [doi: 10.1016/S2352-3018(14)70018-9] [Medline: 26423814]
6. Penedo FJ, Gonzalez JS, Dahn JR, Antoni M, Malow R, Costa P, et al. Personality, quality of life and HAART adherence among men and women living with HIV/AIDS. J Psychosom Res 2003 Mar;54(3):271-278. [doi: 10.1016/s0022-3999(02)00482-8] [Medline: 12614837]
7. Zhu Y, Liu J, Qu B, Hu B, Zhang Y. Relationship between quality of life and unprotected anal intercourse among Chinese men who have sex with men: a cross-sectional study. BMC Public Health 2016 May 10;16:382 [FREE Full text] [doi: 10.1186/s12889-016-3076-z] [Medline: 27165001]
8. Stagl JM, Bouchard LC, Lechner SC, Blomberg BB, Gudenkauf LM, Jutagir DR, et al. Long-term psychological benefits of cognitive-behavioral stress management for women with breast cancer: 11-year follow-up of a randomized controlled trial. Cancer 2015 Jun 01;121(11):1873-1881 [FREE Full text] [doi: 10.1002/cncr.29076] [Medline: 25809235]
9. Antoni MH, Lechner SC, Kazi A, Wimberly SR, Sifre T, Urcuyo KR, et al. How stress management improves quality of life after treatment for breast cancer. J Consult Clin Psychol 2006 Dec;74(6):1143-1152. [doi: 10.1037/0022-006X.74.6.1152] [Medline: 17154743]
10. Berger S, Schad T, von Wyl V, Ehlert U, Zellweger C, Furrer H, et al. Effects of cognitive behavioral stress management on HIV-1 RNA, CD4 cell counts and psychosocial parameters of HIV-infected persons. AIDS 2008 Mar 30;22(6):767-775. [doi: 10.1097/QAD.0b013e3282f511dc] [Medline: 18356607]
11. Penedo FJ, Molton I, Dahn JR, Shen B, Kinsinger D, Traeger L, et al. A randomized clinical trial of group-based cognitive-behavioral stress management in localized prostate cancer: development of stress management skills improves quality of life and benefit finding. Ann Behav Med 2006 Jun;31(3):261-270. [doi: 10.1207/s15324796abm3103_8] [Medline: 16700640]
12. Børøsund E, Ehlers SL, Varsi C, Clark MM, Andrykowski MA, Cvancarova M, et al. Results from a randomized controlled trial testing StressProffen; an application-based stress-management intervention for cancer survivors. Cancer Med 2020 Jun;9(11):3775-3785 [FREE Full text] [doi: 10.1002/cam4.3000] [Medline: 32243717]
13. Price M, Yuen EK, Goetter EM, Herbert JD, Forman EM, Acierno R, et al. mHealth: a mechanism to deliver more accessible, more effective mental health care. Clin Psychol Psychother 2014;21(5):427-436 [FREE Full text] [doi: 10.1002/cpp.1855] [Medline: 23918764]
14. Rathbone AL, Prescott J. The use of mobile apps and SMS messaging as physical and mental health interventions: systematic review. J Med Internet Res 2017 Aug 24;19(8):e295 [FREE Full text] [doi: 10.2196/jmir.7740] [Medline: 28838887]
15. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol 1986 Dec;51(6):1173-1182. [Medline: 3806354]

16. Kraemer HC, Wilson GT, Fairburn CG, Agras WS. Mediators and moderators of treatment effects in randomized clinical trials. Arch Gen Psychiatry 2002 Oct;59(10):877-883. [doi: 10.1001/archpsyc.59.10.877] [Medline: 12365874]
17. Leiberich P, Engeter M, Olbrich E, Rubbert A, Schumacher K, Brieger M, et al. Longitudinal development of distress, coping and quality of life in HIV-positive persons. Psychother Psychosom 1997;66(5):237-247. [doi: 10.1159/000289141] [Medline: 9311027]
18. Zhu M, Guo Y, Li Y, Zeng C, Qiao J, Xu Z, et al. HIV-related stigma and quality of life in people living with HIV and depressive symptoms: indirect effects of positive coping and perceived stress. AIDS Care 2020 Aug;32(8):1030-1035. [doi: 10.1080/09540121.2020.1752890] [Medline: 32290681]
19. Hernández-Ledesma AL, Rodríguez-Méndez AJ, Gallardo-Vidal LS, Trejo-Cruz G, García-Solís P, Dávila-Esquivel FD. Coping strategies and quality of life in Mexican multiple sclerosis patients: physical, psychological and social factors relationship. Mult Scler Relat Disord 2018 Oct;25:122-127. [doi: 10.1016/j.msard.2018.06.001] [Medline: 30071505]
20. Manne SL, Myers-Virtue S, Kashy D, Ozga M, Kissane D, Heckman C, et al. Resilience, positive coping, and quality of life among women newly diagnosed with gynecological cancers. Cancer Nurs 2015;38(5):375-382 [FREE Full text] [doi: 10.1097/NCC.0000000000000215] [Medline: 25521911]
21. Paek M, Ip EH, Levine B, Avis NE. Longitudinal reciprocal relationships between quality of life and coping strategies among women with breast cancer. Ann Behav Med 2016 Oct;50(5):775-783 [FREE Full text] [doi: 10.1007/s12160-016-9803-y] [Medline: 27272631]
22. Folkman S, Moskowitz JT. Coping: pitfalls and promise. Annu Rev Psychol 2004;55:745-774. [doi: 10.1146/annurev.psych.55.090902.141456] [Medline: 14744233]
23. Lutgendorf SK, Antoni MH, Ironson G, Starr K, Costello N, Zuckerman M, et al. Changes in cognitive coping skills and social support during cognitive behavioral stress management intervention and distress outcomes in symptomatic human immunodeficiency virus (HIV)-seropositive gay men. Psychosom Med 1998;60(2):204-214. [doi: 10.1097/00006842-199803000-00017] [Medline: 9560871]
24. Hansen NB, Tarakeshwar N, Ghebremichael M, Zhang H, Kochman A, Sikkema KJ. Longitudinal effects of coping on outcome in a randomized controlled trial of a group intervention for HIV-positive adults with AIDS-related bereavement. Death Stud 2006 Sep;30(7):609-636. [doi: 10.1080/07481180600776002] [Medline: 16865824]
25. Greer JA, Jacobs JM, El-Jawahri A, Nipp RD, Gallagher ER, Pirl WF, et al. Role of patient coping strategies in understanding the effects of early palliative care on quality of life and mood. J Clin Oncol 2018 Jan 01;36(1):53-60 [FREE Full text] [doi: 10.1200/JCO.2017.73.7221] [Medline: 29140772]
26. Gilts CD, Parker PA, Pettaway CA, Cohen L. Psychosocial moderators of presurgical stress management for men undergoing radical prostatectomy. Health Psychol 2013 Dec;32(12):1218-1226 [FREE Full text] [doi: 10.1037/a0030189] [Medline: 23088178]
27. Lemmens LH, Müller VN, Arntz A, Huibers MJ. Mechanisms of change in psychotherapy for depression: an empirical update and evaluation of research aimed at identifying psychological mediators. Clin Psychol Rev 2016 Dec;50:95-107. [doi: 10.1016/j.cpr.2016.09.004] [Medline: 27770716]
28. Guo Y, Hong YA, Qiao J, Xu Z, Zhang H, Zeng C, et al. Run4Love, a mHealth (WeChat-based) intervention to improve mental health of people living with HIV: a randomized controlled trial protocol. BMC Public Health 2018 Dec 26;18(1):793 [FREE Full text] [doi: 10.1186/s12889-018-5693-1] [Medline: 29940921]
29. Guo Y, Hong YA, Cai W, Li L, Hao Y, Qiao J, et al. Effect of a WeChat-based intervention (Run4Love) on depressive symptoms among people living with HIV in China: a randomized controlled trial. J Med Internet Res 2020 Feb 09;22(2):e16715 [FREE Full text] [doi: 10.2196/16715] [Medline: 32044751]
30. Guangzhou population 2020. World Population Review. URL: https://worldpopulationreview.com/world-cities/guangzhou-population [accessed 2020-08-23]
31. WeChat. Wikipedia. URL: https://zh.wikipedia.org/wiki/%E5%BE%AE%E4%BF%A1 [accessed 2020-08-23]
32. Antoni M, Ironson G, Schneiderman N. Cognitive-Behavioral Stress Management for Individuals Living with HIV: Facilitator Guide. Oxford, UK: Oxford University Press; 2007.
33. Zeng Y, Guo Y, Li L, Hong YA, Li Y, Zhu M, et al. Relationship between patient engagement and depressive symptoms among people living with HIV in a mobile health intervention: secondary analysis of a randomized controlled trial. JMIR Mhealth Uhealth 2020 Oct 29;8(10):e20847 [FREE Full text] [doi: 10.2196/20847] [Medline: 33118956]
34. Hsiung P, Fang C, Wu C, Sheng W, Chen S, Wang J, et al. Validation of the WHOQOL-HIV BREF among HIV-infected patients in Taiwan. AIDS Care 2011 Aug;23(8):1035-1042. [doi: 10.1080/09540121.2010.543881] [Medline: 21500023]
35. Zhu Y, Liu J, Qu B. Psychometric properties of the Chinese version of the WHOQOL-HIV BREF to assess quality of life among people living with HIV/AIDS: a cross-sectional study. BMJ Open 2017 Aug 21;7(8):e016382 [FREE Full text] [doi: 10.1136/bmjopen-2017-016382] [Medline: 28827253]
36. Lei H, Cheong CM, Li S, Lu M. The relationship between coping style and internet addiction among mainland Chinese students: a meta-analysis. Psychiatry Res 2018 Dec;270:831-841. [doi: 10.1016/j.psychres.2018.10.079] [Medline: 30551332]
37. Cai Z, Li K, Zhang X. Workplace stressors and coping strategies among Chinese psychiatric nurses. Perspect Psychiatr Care 2008 Oct;44(4):223-231. [doi: 10.1111/j.1744-6163.2008.00181.x] [Medline: 18826460]

38. Xie Y. Reliability and validity of the simplified Coping Style Questionnaire. Chinese J Clin Psychol 1998. [doi: 10.16128/j.cnki.1005-3611.1998.02.018]
39. Cheong J, Mackinnon DP, Khoo ST. Investigation of mediational processes using parallel process latent growth curve modeling. Struct Equ Model 2003 Apr 01;10(2):238 [FREE Full text] [doi: 10.1207/S15328007SEM1002_5] [Medline: 20157639]
40. Bollen KA, Curran PJ. Latent Curve Models: A Structural Equation Perspective. Hoboken, New Jersey, United States: John Wiley & Sons Inc; 2005.
41. Zhao X, Lynch JG, Chen Q. Reconsidering Baron and Kenny: myths and truths about mediation analysis. J Consum Res 2010 Aug 01;37(2):197-206. [doi: 10.1086/651257]
42. Preacher KJ, Hayes AF. SPSS and SAS procedures for estimating indirect effects in simple mediation models. Behav Res Methods Instrum Comput 2004 Nov;36(4):717-731. [doi: 10.3758/bf03206553] [Medline: 15641418]
43. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equat Model Multidiscipl J 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
44. Bentler PM. Fit indexes, lagrange multipliers, constraint changes and incomplete data in structural models. Multivar Behav Res 1990 Apr 01;25(2):163-172. [doi: 10.1207/s15327906mbr2502_3] [Medline: 26794478]
45. Browne MW, Cudeck R. Alternative ways of assessing model fit. Sociol Methods Res 2016 Jun 29;21(2):230-258. [doi: 10.1177/0049124192021002005]
46. Muniz-Terrera G, Robitaille A, Kelly A, Johansson B, Hofer S, Piccinin A. Latent growth models matched to research questions to answer questions about dynamics of change in multiple processes. J Clin Epidemiol 2017 Feb;82:158-166 [FREE Full text] [doi: 10.1016/j.jclinepi.2016.09.001] [Medline: 27639542]
47. Chesney MA, Chambers DB, Taylor JM, Johnson LM, Folkman S. Coping effectiveness training for men living with HIV: results from a randomized clinical trial testing a group-based intervention. Psychosom Med 2003;65(6):1038-1046. [doi: 10.1097/01.psy.0000097344.78697.ed] [Medline: 14645783]
48. Fife BL, Scott LL, Fineberg NS, Zwickl BE. Promoting adaptive coping by persons with HIV disease: evaluation of a patient/partner intervention model. J Assoc Nurses AIDS Care 2008;19(1):75-84 [FREE Full text] [doi: 10.1016/j.jana.2007.11.002] [Medline: 18191771]
49. Xiao Y, Shao Y, Na Z, Zhao W, Wang R, Fang S, et al. A systematic review and meta-analysis of telephone-based therapy targeting depressive symptoms among low-income people living with HIV. AIDS Behav 2021 Mar;25(2):414-426. [doi: 10.1007/s10461-020-02999-8] [Medline: 32809074]
50. Cheng LJ, Kumar PA, Wong SN, Lau Y. Technology-delivered psychotherapeutic interventions in improving depressive symptoms among people with HIV/AIDS: a systematic review and meta-analysis of randomised controlled trials. AIDS Behav 2020 Jun;24(6):1663-1675. [doi: 10.1007/s10461-019-02691-6] [Medline: 31587115]
51. Molassiotis A, Callaghan P, Twinn SF, Lam SW, Chung WY, Li CK. A pilot study of the effects of cognitive-behavioral group therapy and peer support/counseling in decreasing psychologic distress and improving quality of life in Chinese patients with symptomatic HIV disease. AIDS Patient Care STDS 2002 Feb;16(2):83-96. [doi: 10.1089/10872910252806135] [Medline: 11874640]
52. Zhu M, Cai W, Li L, Guo Y, Monroe-Wise A, Li Y, et al. Mediators of intervention effects on depressive symptoms among people living with HIV: secondary analysis of a mobile health randomized controlled trial using latent growth curve modeling. JMIR Mhealth Uhealth 2019 Nov 15;7(11):e15489 [FREE Full text] [doi: 10.2196/15489] [Medline: 31730042]
53. Li Y, Guo Y, Hong YA, Zhu M, Zeng C, Qiao J, et al. Mechanisms and effects of a WeChat-based intervention on suicide among people living with HIV and depression: path model analysis of a randomized controlled trial. J Med Internet Res 2019 Nov 27;21(11):e14729 [FREE Full text] [doi: 10.2196/14729] [Medline: 31774411]

Abbreviations

CBSM: cognitive-behavioral stress management
CFI: confirmatory fit index
LGCM: latent growth curve model
mHealth: mobile health
QOL: quality of life
RCT: randomized controlled trial
RMSEA: root mean square error of approximation
SRMR: standardized root mean square residual
TLI: Tucker-Lewis index
WHOQOL-HIV BREF: World Health Organization Quality of Life HIV short version

Edited by G Eysenbach; submitted 22.11.20; peer-reviewed by JM Simoni, M Ferrer; comments to author 03.06.21; revised version received 05.07.21; accepted 27.12.21; published 17.02.22.

Please cite as:
Zeng Y, Guo Y, Ho RTH, Zhu M, Zeng C, Monroe-Wise A, Li Y, Qiao J, Zhang H, Cai W, Li L, Liu C
Positive Coping as a Mediator of Mobile Health Intervention Effects on Quality of Life Among People Living With HIV: Secondary Analysis of the Randomized Controlled Trial Run4Love
J Med Internet Res 2022;24(2):e25948
URL: https://www.jmir.org/2022/2/e25948
doi:10.2196/25948
PMID:35175209

Review

Digital Health Technologies to Improve Medication Adherence and Treatment Outcomes in Patients With Tuberculosis: Systematic Review of Randomized Controlled Trials

Abdurahman Ridho[1,2], MSc; Sofa D Alfian[1,3], PhD; Job F M van Boven[4,5], PharmD, PhD; Jutti Levita[1,3], Prof Dr; Esin Aki Yalcin[6], Prof Dr; Ly Le[7], Prof Dr; Jan-Willem Alffenaar[8,9,10], Prof Dr; Eelko Hak[11], Prof Dr; Rizky Abdulah[1,3], Prof Dr; Ivan S Pradipta[1,3], PhD

[1]Department of Pharmacology and Clinical Pharmacy, Faculty of Pharmacy, Universitas Padjadjaran, Sumedang, Indonesia

[2]Doctor Program in Pharmacy, Faculty of Pharmacy, Universitas Padjadjaran, Sumedang, Indonesia

[3]Center of Excellence in Higher Education for Pharmaceutical Care Innovation, Universitas Padjadjaran, Sumedang, Indonesia

[4]Department of Clinical Pharmacy and Pharmacology, University Medical Center Groningen, University of Groningen, Groningen, Netherlands

[5]Medication Adherence Expertise Center of the Northern Netherlands, Groningen, Netherlands

[6]Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Ankara University, Ankara, Turkey

[7]Vingroup Big Data Institute, Hanoi, Vietnam

[8]Faculty of Medicine and Health, School of Pharmacy, University of Sydney, Sydney, Australia

[9]Sydney Institute for Infectious Diseases, Sydney, Australia

[10]Westmead Hospital, Sydney, Australia

[11]Pharmacotherapy, Pharmacoepidemiology and Pharmacoeconomics, Research Institute of Pharmacy, University of Groningen, Groningen, Netherlands

Corresponding Author:
Ivan S Pradipta, PhD
Department of Pharmacology and Clinical Pharmacy
Faculty of Pharmacy
Universitas Padjadjaran
Jl. Raya Bandung-Sumedang Km. 21, Jatinangor
Sumedang, 45363
Indonesia
Phone: 62 842 888888 ext 3510
Email: ivanpradipta@unpad.ac.id

Abstract

Background: Nonadherence to medication in tuberculosis (TB) hampers optimal treatment outcomes. Digital health technology (DHT) seems to be a promising approach to managing problems of nonadherence to medication and improving treatment outcomes.

Objective: This paper systematically reviews the effect of DHT in improving medication adherence and treatment outcomes in patients with TB.

Methods: A literature search in PubMed and Cochrane databases was conducted. Randomized controlled trials (RCTs) that analyzed the effect of DHT interventions on medication adherence outcomes (treatment completion, treatment adherence, missed doses, and noncompleted rate) and treatment outcomes (cure rate and smear conversion) were included. Adult patients with either active or latent TB infection were included. The Jadad score was used for evaluating the study quality. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guideline was followed to report study findings.

Results: In all, 16 RCTs were selected from 552 studies found, and 6 types of DHT interventions for TB were identified: 3 RCTs examined video directly observed therapy (VDOT), 1 examined video-observed therapy (VOT), 1 examined an ingestible sensor, 1 examined phone call reminders, 2 examined medication monitor boxes, and 8 examined SMS text message reminders. The outcomes used were treatment adherence, including treatment completion, treatment adherence, missed dose, and noncompleted rate, as well as clinical outcomes, including cure rate and smear conversion. In treatment completion, 4 RCTs (VDOT, VOT, ingestible sensor, SMS reminder) found significant effects, with odds ratios and relative risks (RRs) ranging from 1.10 to 7.69. Treatment adherence was increased in 1 study by SMS reminders (RR 1.05; 95% CI 1.04-1.06), and missed dose was reduced in 1 study by a medication monitor box (mean ratio 0.58; 95% CI 0.42-0.79). In contrast, 3 RCTs of VDOT and 3 RCTs of SMS

reminders did not find significant effects for treatment completion. Moreover, no improvement was found in treatment adherence in 1 RCT of VDOT, missed dose in 1 RCT of SMS reminder, and noncompleted rate in 1 RCT of a monitor box, and 2 RCTs of SMS reminders. For clinical outcomes such as cure rate, 2 RCTs reported that phone calls (RR 1.30; 95% CI 1.07-1.59) and SMS reminders (OR 2.47; 95% CI 1.13-5.43) significantly affected cure rates. However, 3 RCTs found that SMS reminders did not have a significant impact on cure rate or smear conversion.

Conclusions: It was found that DHT interventions can be a promising approach. However, the interventions exhibited variable effects regarding effect direction and the extent of improving TB medication adherence and clinical outcomes. Developing DHT interventions with personalized feedback is required to have a consistent and beneficial effect on medication adherence and outcomes among patients with TB.

(J Med Internet Res 2022;24(2):e33062) doi:10.2196/33062

KEYWORDS

tuberculosis; intervention; eHealth; medication adherence technology; nonadherence; digital health; systematic review; treatment outcomes

Introduction

Tuberculosis (TB), an infectious disease caused by *Mycobacterium tuberculosis* (M. tb), is one of the top 10 most deadly infectious diseases worldwide [1]. The M. tb pathogen can easily spread through air transmission by coughing or sneezing [1]. In 2019, it was estimated that 10 million people globally were infected with TB [1]. Therefore, long-term antibiotic treatment is needed to control TB infection and avoid disease spread.

Treatment for active drug-susceptible TB usually takes at least 6 months, while latent tuberculosis infection (LTBI) can be between 1 and 6 months [2]. The duration of treatment can be longer if the pathogen is resistant to either the first- or second-line of anti-TB medication [3]. Although effective TB medication is available, treatment is prone to nonadherence, resulting in treatment failure [2]. Numerous factors affect medication adherence, such as poor communication between patient and health care provider, socioeconomic status, health care system factors, patients' mental condition, therapy features, and other patient factors [4,5]. Moreover, a high risk of nonadherence to the medication has been reported in patients with LTBI because they do not feel any signs or symptoms of the disease but do experience the side effects of the medication[2].

Previous studies have indicated that nonadherence to TB medication impacts clinical and economic TB outcomes [6]. Notably, the consequences of nonadherence include worsening of the disease [6,7] but also the development and spread of drug-resistant TB [7]. Although poor medication adherence is a widely significant problem in TB treatment, it cannot be managed easily due to the heterogeneity of the underlying factors [5].

Different digital health technologies (DHT) help manage patients with TB by monitoring and supporting their medication adherence [8]. Among them, SMS text message reminders and video directly observed therapy (VDOT) stand out [8]. On the one hand, past evidence indicates that their effect concerning supporting medication adherence among patients with TB is questionable [9,10]. On the other hand, the field of DHTs is evolving rapidly, and some more recent developments may provide more promising approaches in the management of medication adherence. An updated, comprehensive analysis is needed to analyze the effect and potential development of DHTs for managing medication adherence in patients with TB. Therefore, this study systematically reviewed the effect and potential development of various DHTs in improving medication adherence and treatment outcomes in patients with TB.

Methods

Study Design

Randomized controlled studies (RCTs) indexed in the PubMed and Cochrane databases were systematically reviewed following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) guidelines [11,12] (Multimedia Appendix 1).

Data Sources

As a reputable databases in the medical field, PubMed was selected to effectively obtain qualified RCTs. Additionally, the Cochrane Database was included since it provides published and unpublished interventional studies.

Inclusion and Exclusion Criteria

RCTs published in English between March 2002 and January 2020 focusing on interventions using DHT for improving treatment outcomes and medication adherence in patients with TB were included in this review. The time restriction was defined using the leading World Health Organization (WHO) report from 2003 on interventional strategies for improving medication adherence as the starting point [4]. Studies that did not measure the effect of the intervention, had no comparison, applied a different study design, were not original studies (eg, study protocol, review), or were reported as an abstract only were excluded. The population, intervention, comparison, and outcome were predefined to collect articles according to the study's objective.

Population

The study population comprised patients aged 15 years or older with all types of TB (active TB, LTBI, pulmonary TB, extrapulmonary TB, drug-sensitive TB, or drug-resistant TB). TB status needed to be confirmed using laboratory or clinical

testing, such as Mantoux, TB symptoms, chest radiography, interferon gamma release assays, and smear or sputum test; or other examinations, such as polymerase chain reaction or phenotypic drug susceptibility testing.

Interventions and Comparisons

We included studies evaluating the effect of DHTs, such as smartphone apps, video observation, phone reminders, ingestible sensors, SMS reminders, and other digital health interventions, that aimed to improve medication adherence and TB treatment outcomes. All selected studies had a comparison group with patients receiving usual care (mainly traditional directly observed therapy [DOT]) to measure the incremental effect of the intervention.

Outcome Measures

The primary outcome was medication adherence (ie, treatment completion, adherence rate, and missed doses), while the secondary outcome included clinical outcomes (ie, cure rate and smear sputum conversion rate).

Medication adherence could be measured using self-report on the device directly (VDOT) and indirectly (VOT), responding to phone calls and SMS text messages, detecting drug taking through a monitor box, or pill counting. Notably, medication adherence to (respiratory) medicines consists of 3 phases: initiation, implementation, and persistence [13] according to the global TB definition [14] and the ESPACOMP Medication Adherence Reporting Guideline (EMERGE) [15]. In this study, the adherence rate was defined as the rate of anti-TB medication taken by a patient with TB with treatment completion. Following the WHO definition, completion treatment was defined as a patient with TB who completed treatment without evidence of failure (no record of sputum smear or culture test results in the last month of treatment). Hence, the adherence rate was operationally defined as the implementation phase in this study. Moreover, the persistence phase was defined as completion treatment, and noncompletion treatment was deemed to be similar to nonpersistence. Noncompletion treatment was defined as a patient with TB who discontinued the medication before the last defined dose [16,17]. Missed doses were included in the implementation phase and were defined as the percentage of monthly TB medicines missed as measured according to pill count and failure to open the medication box, where the minimum percentage of missed doses was 20% [18].

The clinical outcomes included cure rate, with cure being defined as M. tb culture–positive results at the beginning and negative results in the last month of treatment and on at least 1 previous occasion. The other clinical outcome was sputum conversion, which was defined as the conversion of positive to negative M. tb culture during the treatment [14].

Search Strategies

Specific key terms related to the study population, intervention, comparison, outcome, and design were developed for the specific databases used. The search strategies for the PubMed and Cochrane databases are described in the following sections.

PubMed Database

The search strategy for PubMed was as follows: (“technology”[tw] OR “digital adherence”[tw] OR “mHealth”[tw] OR “mobile health”[tw] OR “mobile app”[tw] OR “mobile apps”[tw] OR “mobile application”[tw]) AND (“medication adherence” [Mesh] OR “adherence”[tw] OR “concordance”[tw] OR “compliance”[tw] OR “nonadherence”[tw] OR “noncompliance”[tw] OR “nonconcordance”[tw]) AND (“tuberculosis/drug therapy”[Mesh] OR “tuberculosis infection”[tw] OR “tb”[tw] OR “active tuberculosis”[tw] OR “latent tuberculosis”[tw] OR “pulmonary tuberculosis”[tw] OR “extrapulmonary tuberculosis”[tw]).

Cochrane Database

The search strategy for Cochrane Database was as follows: (“technology” OR “digital adherence” OR “mHealth” OR “mobile health” OR “mobile app” OR “mobile apps” OR “mobile application”) AND (“medication adherence” [Mesh] OR “adherence” OR “concordance” OR “compliance” OR “nonadherence” OR “noncompliance” OR “nonconcordance”) AND (“tuberculosis/drug therapy” [Mesh] OR “tuberculosis infection” OR “tb” OR “active tuberculosis” OR “latent tuberculosis” OR “pulmonary tuberculosis” OR “extrapulmonary tuberculosis”).

Data Selection, Collection, and Extraction

AR conducted eligibility evaluation based on the title and abstract. The full texts of potentially eligible articles were retrieved and assessed by AR. ISP and SDA conducted further independent verification of the abstract and full-text screening. Any disagreements among the reviewers (AR, ISP, and SDA) were resolved by discussion. Data from the selected articles were extracted by AR and then verified by ISP for relevant information, such as publication year, type of DHT intervention, setting, population, study outcome, and comparison groups.

Summary Measures and Synthesis of Results

In the case of comparable homogenous studies being retrieved, quantitative data synthesis was considered. However, a qualitative narrative review was used in case the data were heterogenous in terms of the population, intervention, comparisons, or outcomes. Several point estimates were considered for the data analysis, such as the mean ratio for continuous outcome data and the relative risk (RR) and odds ratio (OR) for dichotomous outcome data with a 95% CI.

Quality Assessment of the Included Articles

As all included articles were RCTs, the Jadad score was used to assess the individual articles [19]. The Jadad score has 3 assessment domains: randomization, blinding method, and participant withdrawal with a minimum score of 1 (poor quality) and a maximum score of 5 (good quality).

Results

Study Selection

Our literature search initially identified 552 articles. After removal of duplicates, the screening of titles and abstracts

yielded 20 relevant articles. The full-text screening process resulted in a final study sample of 16 articles. Variations in the study population, comparisons, and outcomes across the included studies were identified; therefore, quantitative analysis could not be conducted in this study considering the heterogeneity. The flow diagram and the screening process are provided in Figure 1.

The 16 included studies were conducted in different countries: England [20], USA [21-23], Taiwan [24], China [17,25], Pakistan [26,27], South Africa [9], Cameroon [28], Canada [10], Argentina [29], Sudan [17], Thailand [30], and Haiti [31].

The median sample size of the source population was 259 participants, ranging from 37 participants [29] to 1110 participants [24]. Most studies targeted either patients alone (n=12) or patients and health care professionals (n=4). Several study designs were used in the included articles: 11 RCTs with a 2-arm design [10,17,21-24,26,27,29-31], 4 with a 2-arm cluster design [9,20,25,28], and 1 with a 4-arm cluster design [17].

Across the studies, multiple various interventions were assessed, including 3 assessing VDOT [21,22,24], 1 VOT [20], 1 phone call reminders [30], 2 medication monitor boxes [18,31], 1 ingestible sensors [23], and 8 SMS reminders [9,10,17,25-29]. The characteristics of all included articles are indicated in Table 1.

Figure 1. Flow diagram of the article selection process.

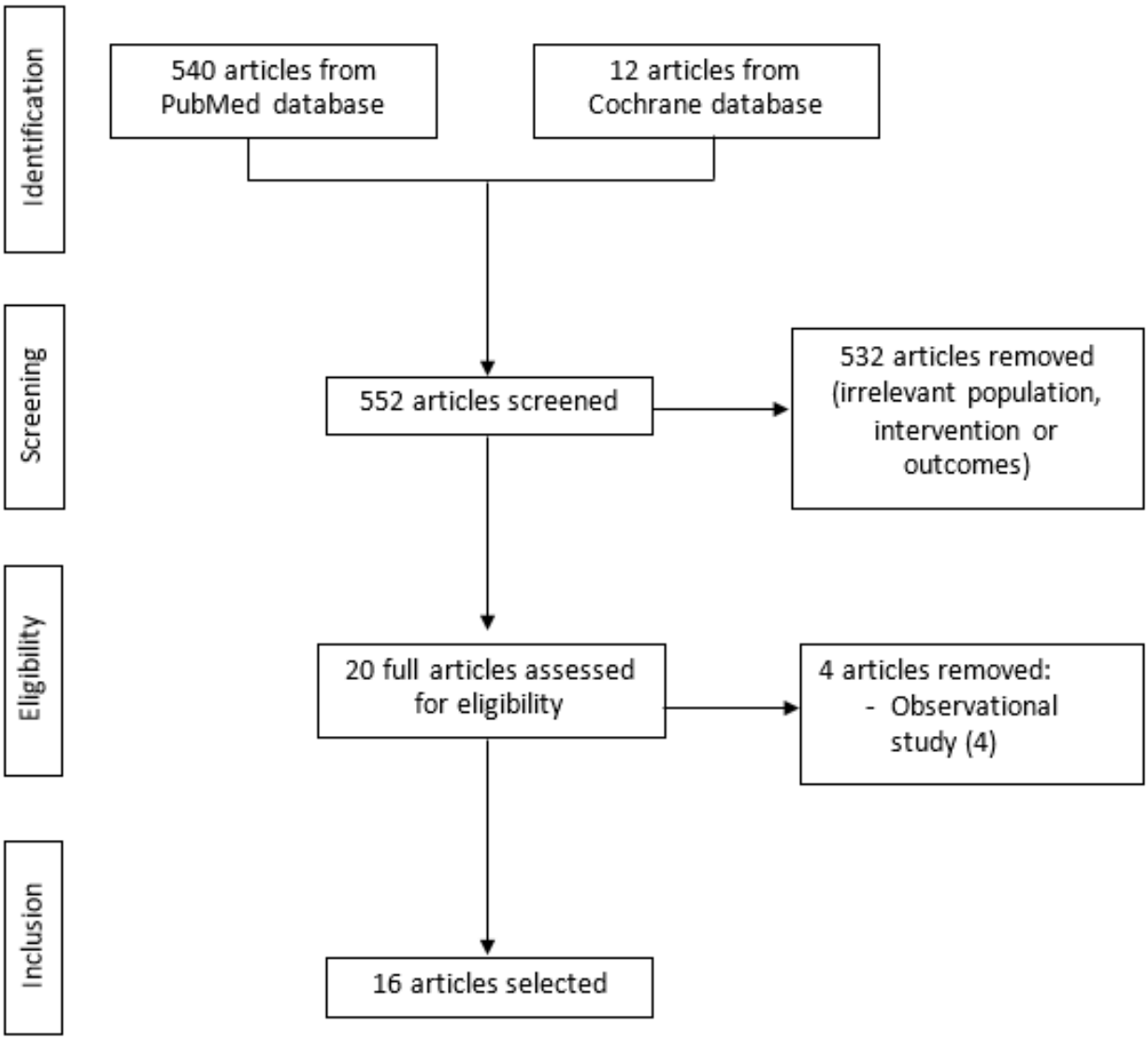


Table 1. Characteristics of the included articles.

No.	Authors, publication year	Target	Study period	Setting	Intervention	Participants in the intervention and comparison group, n	Study outcomes	Outcome measurements
1	Story et al, 2019 [20]	Patients	Sept 2014 to Oct 2016	22 clinics in England (UK)	VOT[a] intervention in which patients record and send videos of every dose ingested and adverse medicine events using the smart app[b]	112 and 114	Treatment completion	Scheduled treatment observation
2	Chuck et al, 2016 [21]	Patients	Sept 2013 to Sept 2014	TB[c] clinics in New York City, USA	Patients swallowing a pill in front of a camera on their schedule on VDOT[d] calls[b]	49 and 267	Treatment completion	Scheduled VDOT[d] session
3	Lam et al, 2018 [22]	Patients	Feb to Oct 2015	4 New York City health department TB clinics, New York City, USA	Patients taking medicine observed by a VDOT worker[b]	50 and 302	Treatment completion	Scheduled VDOT session
4	Chen et al, 2020 [24]	Patients and HCP[e]	Jan 2014 to Dec 2017	Health facilities in Taipei, Taiwan	Patients taking medicine under 2-way video calls[b]	80 and 160	Treatment adherence	Scheduled VDOT session
5	Kunawararak et al, 2011 [30]	Patients	Apr 2008 to Dec 2009	Public hospitals in 7 provinces of Northern Thailand	Patients receiving phone call reminders to take their medication[b]	30 and 30	Cure rate	Negative culture result at the end of treatment
6	Liu et al, 2015 [18]	Patients	June 2011 to Mar 2012	36 districts within the province of Heilongjiang, Jiangsu, Hunan, and Chongqing, China	Patients receiving a reminder via SMS, medication monitor box, or their medication[b]	3069 and 1104	Missed doses	Pill count and failure to open the medication monitor box
7	Moulding & Caymittes, 2002 [31]	Patients	July 1983 to Nov 1985	Clinic at Port-Au-Prince, Haiti	Medication monitor and counseling in which the medication monitor is based on moving a minute piece of uranium along a strip of photographic film to record the interval between the removal of each tablet of medication[f]	64 and 127	Noncompleted rate	Counting the number of dots on the film strip
8	Browne et al, 2019 [23]	Patients	Oct 2013 to Jan2017	San Diego and Orange County divisions of TB Control and Refugee Health, USA	Patients given the WOT[g] IS-Rifamate (ingestible sensor)[i]	41 and 20	Treatment completion	The number of doses confirmed
9	Fang et al, 2017 [25]	Patients	Dec 2014 to Dec 2015	6 districts from Anhui province, China	An SMS reminder sent once per day to remind patients for taking the medicine, reexamining the physical condition, and improving knowledge[b]	160 and 190	Treatment completion and missed doses	Pill counting
10	Mohammed et al, 2016 [26]	Patients and HCP[e]	Mar 2011 to Feb 2014	Public and private sector tuberculosis clinics in Karachi, Pakistan	Daily SMS reminders sent to participants who are asked to respond through SMS or missed calls after taking their medication[b]	1110 and 1097	Treatment completion and smear conversion	Pill counting and smear conversion at the end of treatment

No.	Authors, publication year	Target	Study period	Setting	Intervention	Participants in the intervention and comparison group, n	Study outcomes	Outcome measurements
11	Belknap et al, 2017 [9]	Patients	Sept 2012 to Apr 2014	Outpatient tuberculosis clinics in the USA, Spain, Hong Kong, and South Africa	SAT[h] used without reminder and SAT used with weekly SMS reminder[b]	315 and 328	Treatment completion	Pill counting and self-report
12	Bediang et al, 2018 [28]	Patients	Feb 2013 to Apr 2014	Treatment and Diagnostic Centres of Yaoundé, Cameroon	Daily SMS reminder for 6 months[b]	137 and 142	Treatment completion and cure rate	Pill counting and negative culture result at the end of treatment
13	Johnston et al, 2018 [10]	Patients and HCP	June 2012 to Sept 2015	TB clinics in British Columbia, Canada	2-way weekly SMS reminder to patients who need to respond within 48 hours[i]	170 and 188	Treatment completion	Pill counting and self-report
14	Iribarren et al, 2013 [29]	Patients and HCP	Nov 2011 to June 2012	Clinic located within Health Region V in the province of Buenos Aires, Argentina	Patients receiving SMS reminders every day and being required to text after medication administration[i]	18 and 19	Treatment completion and treatment adherence	Pill counting and self-report
15	Ali & Martin, 2019 [17]	Patients	May 2017 to Mar 2018	8 TB treatment units in Khartoum Province, Khartoum State, Sudan	Patients receiving standard of care, with the additional text messages every 48 hours during the first 2 months and weekly thereafter until the end of treatment[i]	74 and 74	Cured rate and noncompleted rate	Negative culture result at the end of treatment and pill counting
16	Farooqi et al, 2017 [27]	Patients	June 2014 to June 2015	TB clinics of Khyber Teaching Hospital Peshawar and Emergency Satellite Hospital Nahaqi, Pakistan	Patients receiving daily SMS reminders and DOT[b]	74 and 74	Cure rate and noncompleted rate	Negative culture result at the end of treatment and pill counting

[a]VOT: video-observed treatment.

[b]The comparison group is the directly observed treatment (DOT).

[c]TB: tuberculosis.

[d]VDOT: video directly observed treatment.

[e]HCP: health care provider.

[f]The comparison group is the simple container.

[g]WOT: wirelessly observed therapy.

[h]SAT: self-administered therapy.

[i]The comparison group is the standard of care.

Outcomes

The medication adherence-related outcomes reported across the 16 RCTs were treatment completion in 10 (63%) studies [9,10,20-23,25,26,28,29], treatment adherence in 2 (13%) [24,29], missed doses in 2 (13%) [18,25], and noncompleted rate in 3 (19%) [17,27,31]. Note that some RCTs reported more than 1 adherence outcome. Clinical outcomes included cure rate in 4 RCTs (25%) [17,27,28,30] and smear conversion in 1 RCT [26] (6%).

Adherence Outcomes: Positive Findings

Four RCTs showed that treatment completion was significantly and positively impacted by a DHT intervention, with effect sizes varying by type of intervention. Notably, VDOT or VOT was found to be more effective compared to DOT in 2 studies from the UK and USA (UK: OR 2.52; 95% CI 1.17-5.54 [20]; USA: RR 1.36; 95% CI 1.19-1.55 [22]). Furthermore, studies on ingestible sensors (OR 7.69; 95% CI 4.51-14.48) [23] and 1 of the SMS reminder intervention studies (RR 1.1; 95% CI 1.04–1.18) [25] reported positive effects on treatment completion.

The outcome of treatment adherence was increased in 1 study that evaluated SMS reminders (RR 1.05; 95% CI 1.04-1.06) [29], and missed doses was reduced in 1 study that used a medication monitor box (mean ratio 0.58; 95% CI 0.42-0.79) [18] in patients with TB.

Adherence Outcomes: Negative Findings

Beyond the trials that reported the positive impact of DHT interventions on adherence outcomes, several studies that did not report any significant effect on medication adherence outcomes were also identified. Three RCTs did not find any significant effect on treatment completion from VDOT (RR 0.99, 95% CI 0.93-1.05 [21]; RR 1.00, 95% CI 0.79-1.26 [26]; RR 0.87, 95% CI 0.81-0.94 [9]), and another 3 indicated a lack of evidence for a positive effect of SMS reminders (OR 1.45, 95% CI 0.81-2.56 [28]; RR 0.97, 95% CI 0.88-1.07 [10]; RR 0.99, 95% CI 0.93-1.05 [29]).

Additionally, treatment adherence was not impacted by VDOT in 1 RCT (RR 1.08; 95% CI 0.89-1.32 [24]), and neither were missed doses by SMS reminders (RR 0.40; 95% CI 0.11-1.50 [25]), or the noncompleted rate by a monitor box (RR 0.55; 95% CI 0.21-1.42 [31]), or SMS reminders (OR 1.67, 95% CI 0.52-5.37 [17]; RR 0.76, 95% CI 0.18-3.28 [27]).

Clinical Outcomes: Positive Findings

In 2 studies, the clinical outcome of cure rate was positively affected by DHT [17,30]. It was observed that cure rate was increased through SMS reminders (OR 2.47, 95% CI 1.13-5.43 [17]) and phone call reminders (RR 1.30, 95% CI 1.07-1.59 [30]) in patients with TB.

Clinical Outcomes: Negative Findings

Three studies did not find there to be a significant effect of DHT on clinical outcomes, with 2 RCTs indicating no effects of SMS reminders on cure rate (OR 1.06;,95% CI 0.65-1.73 [28]; RR 1.05, 95% CI 0.62-1.76 [27]) and 1 study indicating no effect of SMS reminders on smear conversion (RR 1.00, 95% CI 0.90-1.12) [26]. The effects of interventions are summarized in Table 2.

Table 2. Effects of digital health technology interventions on medication adherence and clinical tuberculosis outcomes.

No.	First author, year of publication	Intervention[a]	Medication adherence outcome (95% CI)				Clinical outcome (95% CI)	
			Treatment completion	Treatment adherence	Missed doses	Noncompleted rate	Cure rate	Smear sputum conversion
1	Story et al, 2019 [20]	VOT[b]	OR[c]: 2.52 (1.17-5.54)	N/A[d]	N/A	N/A	N/A	N/A
2	Chuck et al, 2016 [21]	VDOT[e]	RR[f]: 0.99 (0.93-1.05)	N/A	N/A	N/A	N/A	N/A
3	Lam et al, 2018 [22]	VDOT	RR: 1.36 (1.19-1.55)	N/A	N/A	N/A	N/A	N/A
4	Chen et al, 2020 [24]	VDOT	N/A	RR: 1.08 (0.89-1.32)	N/A	N/A	N/A	N/A
5	Kunawararak et a., 2011 [30]	Phone call reminder	N/A	N/A	N/A	N/A	RR: 1.30 (1.07-1.59)	N/A
6	Liu et al, 2015 [18]	SMS, medication monitor box, and combined	N/A	N/A	MR[g]: 0.94 (0.71-1.24) MR: 0.58 (0.42-0.79) MR: 0.49 (0.27-0.88)	N/A	N/A	N/A
7	Moulding & Caymittes, 2002 [31]	Medication monitor with counseling	N/A	N/A	N/A	RR: 0.55 (0.21-1.42)	N/A	N/A
8	Browne et al, 2019 [23]	Ingestible sensor	OR: 7.69 (4.51-14.48)	N/A	N/A	N/A	N/A	N/A
9	Fang et al, 2017 [25]	SMS reminder	RR: 1.10 (1.04-1.18)	N/A	RR: 0.40 (0.11-1.50)	N/A	N/A	N/A
10	Mohammed et al, 2016 [26]	SMS reminder	RR: 1.00 (0.79–1.26)	N/A	N/A	N/A	N/A	RR: 1.00 (0.90–1.12)
11	Belknap et al, 2017 [9]	SMS reminder	RR: 0.87 (0.81-0.94)	N/A	N/A	N/A	N/A	N/A
12	Bediang et al., 2018 [28]	SMS reminder	OR: 1.45 (0.81-2.56)	N/A	N/A	N/A	OR: 1.06 (0.65-1.73)	N/A
13	Johnston et al, 2018 [10]	SMS reminder	RR: 0.97 (0.88-1.07)	N/A	N/A	N/A	N/A	N/A
14	Iribarren et al, 2013 [29]	SMS reminder	RR: 0.99 (0.93-1.05)	RR: 1.05 (1.04-1.06)	N/A	N/A	N/A	N/A
15	Ali & Martin, 2019 [17]	SMS reminder	N/A	N/A	N/A	OR 1.67 (0.52-5.37)	OR: 2.47 (1.13-5.43)	N/A
16	Farooqi et al, 2017 [27]	SMS reminder	N/A	N/A	N/A	RR: 0.76 (0.18-3.28)	RR: 1.05 (0.62-1.76)	N/A

[a]The comparison group is provided in Table 1.

[b]VOT: video-observed therapy.

[c]OR: odds ratio.

[d]N/A: not applicable.

[e]VDOT: video directly observed therapy.

[f]RR: relative risk.

[g]MR: mean ratio.

Quality Assessment of the Included Articles

None of the included studies had the maximum Jadad score (5 points). Eleven studies (69%) had a relatively high score (3 points) since they appropriately described the randomization method and clearly illustrated the withdrawals of participants [9,10,18,20,23,25-29,31]. Furthermore, 5 of the 16 studies (31%) had lower Jadad scores (2 points) due to the absence of a description of the randomization procedure [17,21,22,24,30].

The risk of bias assessment is indicated in (Multimedia Appendix 2).

Data Synthesis

As shown in Table 1, considerable variations in the study population, comparisons, and outcomes across the included studies were identified. Therefore, the quantitative analysis could not be conducted considering the heterogeneity observed.

Discussion

Main Findings

This systematic review indicated that DHT interventions, including VOT, VDOT, medication monitor boxes, ingestible sensors, and SMS reminders, could effectively improve medication adherence in patients with TB. However, half of the 4 V(D)OT and over half of the 8 SMS reminder studies showed no significant effect on adherence outcomes, resulting in inconclusive evidence regarding their effectiveness with respect to the extent of improving medication adherence. Regarding reported clinical outcomes, 1 study with phone call reminders and 1 of the 3 studies with SMS reminders showed positive effects on TB cure rates, and no DHT exhibited effects on sputum conversion rates.

Among the primary DHT interventions, VDOT and VOT are the interventions with the potential for improving TB medication adherence. With video-observed treatment, medication adherence can be directly recorded (VDOT) or indirectly (VOT) while a patient takes the medication. Hence, the health care provider can monitor the medication intake process remotely. However, this technology needs adequate smartphone specifications (eg, picture resolution, memory), good connectivity, and user ability. Smartphones with good connectivity are required because each video should be sent in an adequate video resolution and fairly large file size for proper drug monitoring. The benefit of video treatment is that patients' and health care workers' travel time can be significantly decreased compared to traditional DOT [22]. As an intervention, VDOT or VOT is preferred by patients with TB because it is easier to accept, cheaper, and more effective than is standard DOT [20]. Moreover, the implementation of VDOT or VOT has been associated with treatment completion [20-22] and treatment adherence [24] in patients with TB. Our review found that 2 VDOT studies significantly enhanced treatment completion compared to DOT [20,22], while 2 other studies demonstrated no difference between VDOT and standards of care [21,24]. The challenges concerning its implementation are the availability of (smart) phones and the network-related interruption of audio-video connections during VDOT sessions. However, the widespread use of smartphones will encourage providers to improve network quality. Therefore, VDOT or VOT can be a promising approach to overcoming medication adherence problems [32], especially in countries with high smartphone use [33,34]. The average availability of smartphones in the top-30 high-burden TB countries is 93.98 per 100 people, ranging from 14.9 in North Korea to 185.69 in Montenegro [35]. The data highlighted the potential use of DHT interventions based on smartphone intervention.

Beyond VDOT, SMS reminders were one of the most frequently studied DHT interventions. Phone call reminders are used to remind patients directly to take their medicine. The technology used in this approach is simpler than that of VDOT or VOT because the phone call reminder does not need high device specifications and patient ability; thus, patients can easily use it. One study in Thailand that evaluated 2 models for TB control, single DOT and DOT with phone call reminders, showed that phone call reminders can increase the cure rate [29]. The reminders supported the patients during their treatment because they did not feel alone or socially isolated. Another type of reminder-based intervention is the medication monitor box. This device can also record the history of medicine usage and can be used for adherence monitoring by health care providers at a distance [18]. The medication reminder feature can overcome forgetfulness in patients with TB. A study conducted in 36 districts in China using a medication monitor box with a reminder exhibited a lower missed-dose rate than did a medication monitor box without a reminder [17]. Another RCT conducted in Haiti reported that the medication monitor box without reminders did not effectively reduce the noncompleted rate among patients with TB [31]. The study highlighted that the reminder feature attached in the medication box could affect the completion rate.

Another DHT is the ingestible sensor. The sensor is attached to the medicine and sends a signal to the operator when the medicine is ingested. If the patient is not taking the medication, the system will automatically send a reminder urging them to take medicine immediately. With this ingestible sensor, the potential adverse effects should be considered since the sensor is attached to the patient. An RCT conducted in the USA on patients with advanced-phase TB indicated that ingestible sensors could increase treatment completion with an accuracy of 99.3%. However, rash and pruritus were reported as side effects in up to 10% of the participants [23]. It would be worthwhile to conduct a cost-effectiveness analysis to assess the increased treatment completion at the additional cost of the ingestible sensors and the side effects.

SMS reminders can be considered and combined with other interventions to improve medication adherence in patients with TB. SMS reminders are sent periodically to encourage patients to take their medicine. This feature does not require high smartphone specifications. Indeed, 3 of the 8 RCT studies conducted in 6 countries indicated that the SMS reminder could increase treatment completion [25], treatment adherence [29], and cure rate [29]. The SMS reminder can be applied widely because of its ease of access, low cost, and ready acceptance by patients. Like the medication monitor box, the SMS reminder plays an essential role in improving patients' medication schedules [25]. However, 5 studies found no benefit compared to the standard of care [9,10,17,26,28]. This implies that careful consideration should be made before implementing this intervention. Implementation under operational research conditions would be ideal as data would become available to evaluate the benefit of the SMS reminder service.

Some studies have indicated that DHT interventions could improve treatment outcomes and patient adherence. However, some studies exhibited no difference compared to the standard

of care [9,10,17,21,24,26,28,31]. The variable effects of DHT interventions can be explained using the heterogeneity of the population, comparison, and outcome definitions used across the included studies. For instance, although the same type of intervention and outcome was applied, the use of various comparisons (DOT and standard care without DOT) [10,25] makes it challenging to aggregate results. Notably, the features of the population should be taken into consideration when interpreting the intervention effect on adherence. As nonadherence can be explained using multiple factors (eg, social, economic, and behavioral factors) [4], it is essential to characterize the individual problem of nonadherence when conducting the intervention. Hence, a one-size-fits-all intervention to resolve nonadherence is unlikely to be successful. Therefore, a tailored and targeted intervention for improving medication adherence is essential to resolving nonadherence.

Implications and Future Directions

Given the contradicting study outcomes, more research and developments regarding the role of DHTs in medication adherence management for patients with TB are needed. These developments should focus on the (technical) drug monitoring aspects and the intervention(s) that can effectively improve medication adherence and treatment outcome. Considering the multiple factors underlying medication nonadherence in patients with TB (eg, social, economic, geographic, facility, and behavioral factors) [4], a tailored and targeted intervention for improving medication adherence is essential. Therefore, a DHT with screening and monitoring for nonadherent patients can be further enhanced by considering the individual problems patients encounter. Subsequently, these data can further support health care providers in effectively delivering personalized interventions for improving medication adherence in patients with TB. This may also require further training of health care providers in effective communication strategies.

Most DHT-based interventions in this study were assessed under trial conditions that may not represent the real-world condition for implementation. Numerous factors may drive the effect of a DHT's intervention in the real world: population density, facilities, transportation, smartphone network coverage, health care systems, human resources, costs, individual characteristics, and integration of the intervention with the national TB program [36]. This underlines the fact that implementing a particular intervention will have very different meanings depending on the local context. Therefore, understanding the local context is critical for successful DHT intervention implementation in real-world settings. Additionally, governmental policy and proper reimbursement are needed to support and regulate DHTs for successful implementation. Further implementation studies are needed to evaluate the maturity of interventions in the real world to scale up DHT interventions beyond trial settings.

Strengths and Limitations

This study's strength is that a comprehensive and up-to-date overview of existing DHTs, ranging from their use, device specifications, and efficacy that can be used to improve the problem of medication adherence and treatment outcome among patients with TB, was provided. However, although all efforts were made to provide a robust analysis, several limitations should be acknowledged: although unpublished RCT studies were covered in the Cochrane Database, potential publication bias may exist in this study due to limited inclusion of gray literature, non–English language studies, and studies indexed in other databases; a meta-analytical analysis could not be conducted due to the heterogeneity across the included studies; as the focus was on RCT studies—known as the gold standard in analyzing the effects of interventions—well-designed pre-and postintervention studies were excluded in this study.

Conclusions

It was found that DHT can be a promising approach in improving medication adherence and treatment outcomes among patients with TB despite the variable intervention effects that were discovered. Considering individual factors of nonadherence to medication among patients with TB, developing DHT interventions with personalized feedback is required to have a consistent and beneficial effect on medication adherence and treatment outcomes among patients with TB. Further implementation studies are needed to evaluate the maturity and scale-up of the interventions in the real world.

Acknowledgments

The authors wish to thank Prof. Federico Gago, University of Alcala, Madrid, Spain, for his support on this study and suggestions on preparing the manuscript. This study was funded by the Ministry of Education, Culture, Research and Technology of the Republic of Indonesia, under the Southeast Asia–Europe Joint Funding for Research and Innovation 2019 (grant #9/AMD/E1/KP.PTNBH/2020) and the Scientific and Technological Research Council of Turkey ARDEB International (JFS2017; grant #119N553).

Conflicts of Interest

None declared.

Multimedia Appendix 1
The PRISMA checklist.
[DOCX File , 24 KB - jmir_v24i2e33062_app1.docx]

Multimedia Appendix 2

The Jadad score table.
[DOCX File , 15 KB - jmir_v24i2e33062_app2.docx]

References

1. World Health Organization. Global Tuberculosis Report 2020. Geneva: World Health Organization; Oct 15, 2020.
2. World Health Organization. Guidelines for Treatment of Medicines-Susceptible Tuberculosis and Patient Care 2017 Update. Geneva: World Health Organization; Apr 24, 2017.
3. World Health Organization. Consolidated Guidelines on Drug-Resistant Tuberculosis Treatment. Geneva: World health Organization. Geneva: World health Organization; Mar 20, 2019.
4. World Health Organization. Adherence to Long-Term Therapies. Geneva: World Health Organization. Geneva: World Health Organization; 2003.
5. Pradipta I, Idrus L, Probandari A, Lestari B, Diantini A, Alffenaar J, et al. Barriers and strategies to successful tuberculosis treatment in a high-burden tuberculosis setting: a qualitative study from the patient's perspective. BMC Public Health 2021 Oct 21;21(1):1903 [FREE Full text] [doi: 10.1186/s12889-021-12005-y] [Medline: 34670527]
6. Chimeh RA, Gafar F, Pradipta IS, Akkerman OW, Hak E, Alffenaar JC, et al. Clinical and economic impact of medication non-adherence in drug-susceptible tuberculosis: a systematic review. Int J Tuberc Lung Dis 2020 Aug 01;24(8):811-819. [doi: 10.5588/ijtld.19.0754] [Medline: 32912386]
7. Pradipta IS, Forsman LD, Bruchfeld J, Hak E, Alffenaar J. Risk factors of multidrug-resistant tuberculosis: A global systematic review and meta-analysis. J Infect 2018 Dec;77(6):469-478. [doi: 10.1016/j.jinf.2018.10.004] [Medline: 30339803]
8. World Health Organization. Handbook for the Use of Digital Technologies to Support Tuberculosis Medication Adherence. Geneva: World Health Organization; 2017.
9. Belknap R, Holland D, Feng P, Millet J, Caylà JA, Martinson NA, TB Trials Consortium iAdhere Study Team. Self-administered versus directly observed once-weekly isoniazid and rifapentine treatment of latent tuberculosis infection: a randomized trial. Ann Intern Med 2017 Nov 21;167(10):689-697 [FREE Full text] [doi: 10.7326/M17-1150] [Medline: 29114781]
10. Johnston JC, van der Kop ML, Smillie K, Ogilvie G, Marra F, Sadatsafavi M, et al. The effect of text messaging on latent tuberculosis treatment adherence: a randomised controlled trial. Eur Respir J 2018 Feb 07;51(2):1701488 [FREE Full text] [doi: 10.1183/13993003.01488-2017] [Medline: 29437940]
11. Pradipta I, Houtsma D, van Boven JFM, Alffenaar J, Hak E. Interventions to improve medication adherence in tuberculosis patients: a systematic review of randomized controlled studies. NPJ Prim Care Respir Med 2020 May 11;30(1):21 [FREE Full text] [doi: 10.1038/s41533-020-0179-x] [Medline: 32393736]
12. Moher D, Liberati A, Tetzlaff J, Altman D, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol 2009 Oct;62(10):1006-1012. [doi: 10.1016/j.jclinepi.2009.06.005] [Medline: 19631508]
13. Vrijens B, Dima AL, Van GE, van BJFM, Eakin MN, Foster JM, et al. What we mean when we talk about adherence in respiratory medicine. J Allergy Clin Immunol Pract 2016 Sep;4(5):802-812 [FREE Full text] [doi: 10.1016/j.jaip.2016.05.019] [Medline: 27587314]
14. Definitions and Reporting Framework for Tuberculosis: 2013 Revision (Updated December 2014 and January 2020). Geneva: World Health Organization; 2013.
15. De Geest S, Zullig LL, Dunbar-Jacob J, Helmy R, Hughes DA, Wilson IB, et al. ESPACOMP Medication Adherence Reporting Guideline (EMERGE). Ann Intern Med 2018 Jul 03;169(1):30-35 [FREE Full text] [doi: 10.7326/M18-0543] [Medline: 29946690]
16. Podewils LJ, Gler MTS, Quelapio MI, Chen MP. Patterns of treatment interruption among patients with multidrug-resistant TB (MDR TB) and association with interim and final treatment outcomes. PLoS One 2013 Jul 29;8(7):e70064 [FREE Full text] [doi: 10.1371/journal.pone.0070064] [Medline: 23922904]
17. Ali AOA, Prins MH. Mobile health to improve adherence to tuberculosis treatment in Khartoum state, Sudan. J Public Health Afr 2019 Dec 31;10(2):1101 [FREE Full text] [doi: 10.4081/jphia.2019.1101] [Medline: 32257081]
18. Liu X, Lewis JJ, Zhang H, Lu W, Zhang S, Zheng G, et al. Effectiveness of electronic reminders to improve medication adherence in tuberculosis patients: a cluster-randomised trial. PLoS Med 2015 Sep;12(9):e1001876 [FREE Full text] [doi: 10.1371/journal.pmed.1001876] [Medline: 26372470]
19. Jadad AR, Moore R, Carroll D, Jenkinson C, Reynolds DM, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials 1996 Feb;17(1):1-12. [doi: 10.1016/0197-2456(95)00134-4] [Medline: 8721797]
20. Story A, Aldridge RW, Smith CM, Garber E, Hall J, Ferenando G, et al. Smartphone-enabled video-observed versus directly observed treatment for tuberculosis: a multicentre, analyst-blinded, randomised, controlled superiority trial. Lancet 2019 Mar 23;393(10177):1216-1224 [FREE Full text] [doi: 10.1016/S0140-6736(18)32993-3] [Medline: 30799062]
21. Chuck C, Robinson E, Macaraig M, Alexander M, Burzynski J. Enhancing management of tuberculosis treatment with video directly observed therapy in New York City. Int J Tuberc Lung Dis 2016 May 01;20(5):588-593. [doi: 10.5588/ijtld.15.0738] [Medline: 27084810]

22. Lam CK, McGinnis Pilote K, Haque A, Burzynski J, Chuck C, Macaraig M. Using video technology to increase treatment completion for patients with latent tuberculosis infection on 3-month isoniazid and rifapentine: An implementation study. J Med Internet Res 2018 Nov 20;20(11):e287 [FREE Full text] [doi: 10.2196/jmir.9825] [Medline: 30459146]
23. Browne SH, Umlauf A, Tucker AJ, Low J, Moser K, Gonzalez Garcia J, et al. Wirelessly observed therapy compared to directly observed therapy to confirm and support tuberculosis treatment adherence: A randomized controlled trial. PLoS Med 2019 Oct 4;16(10):e1002891 [FREE Full text] [doi: 10.1371/journal.pmed.1002891] [Medline: 31584944]
24. Chen S, Wang I, Hsu H, Huang C, Liu Y, Putri DU, et al. Advantage in privacy protection by using synchronous video observed treatment enhances treatment adherence among patients with latent tuberculosis infection. J Infect Public Health 2020 Sep;13(9):1354-1359 [FREE Full text] [doi: 10.1016/j.jiph.2020.03.013] [Medline: 32376234]
25. Fang X, Guan S, Tang L, Tao F, Zou Z, Wang J, et al. Effect of short message service on management of pulmonary tuberculosis patients in Anhui Province, China: A prospective, randomized, controlled study. Med Sci Monit 2017 May 23;23:2465-2469 [FREE Full text] [doi: 10.12659/msm.904957] [Medline: 28534476]
26. Mohammed S, Glennerster R, Khan AJ. Impact of a daily SMS medication reminder system on tuberculosis treatment outcomes: A randomized controlled trial. PLoS One 2016 Nov 1;11(11):e0162944 [FREE Full text] [doi: 10.1371/journal.pone.0162944] [Medline: 27802283]
27. Farooqi RJ, Ashraf S, Zaman M. The role of mobile SMS-reminders in improving drugs compliance in patients receiving anti-TB treatment from DOTS program. Journal of Postgraduate Medical Institute 2017 May 27;31(2):156-162 [FREE Full text]
28. Bediang G, Stoll B, Elia N, Abena J, Geissbuhler A. SMS reminders to improve adherence and cure of tuberculosis patients in Cameroon (TB-SMS Cameroon): a randomised controlled trial. BMC Public Health 2018 May 02;18(1):583 [FREE Full text] [doi: 10.1186/s12889-018-5502-x] [Medline: 29720146]
29. Iribarren S, Beck S, Pearce PF, Chirico C, Etchevarria M, Cardinale D, et al. TextTB: a mixed method pilot study evaluating acceptance, feasibility, and exploring initial efficacy of a text messaging intervention to support TB treatment adherence. Tuberc Res Treat 2013;2013:349394-349312 [FREE Full text] [doi: 10.1155/2013/349394] [Medline: 24455238]
30. Kunawararak P, Pongpanich S, Chantawong S, Pokaew P, Traisathit P, Srithanaviboonchai K, et al. Tuberculosis treatment with mobile-phone medication reminders in northern Thailand. Southeast Asian J Trop Med Public Health 2011 Nov;42(6):1444-1451. [Medline: 22299414]
31. Moulding TS, Caymittes M. Managing medication compliance of tuberculosis patients in Haiti with medication monitors. Int J Tuberc Lung Dis 2002 Apr;6(4):313-319. [Medline: 11936740]
32. Guo X, Yang Y, Takiff HE, Zhu M, Ma J, Zhong T, et al. A comprehensive app that improves tuberculosis treatment management through video-observed therapy: usability study. JMIR Mhealth Uhealth 2020 Jul 31;8(7):e17658 [FREE Full text] [doi: 10.2196/17658] [Medline: 32735222]
33. Rabinovich L, Molton JS, Ooi WT, Paton NI, Batra S, Yoong J. Perceptions and acceptability of digital interventions among tuberculosis patients in Cambodia: qualitative study of video-based directly observed therapy. J Med Internet Res 2020 Jul 27;22(7):e16856 [FREE Full text] [doi: 10.2196/16856] [Medline: 32716309]
34. Holzman SB, Atre S, Sahasrabudhe T, Ambike S, Jagtap D, Sayyad Y, et al. Use of Smartphone-Based Video Directly Observed Therapy (vDOT) in Tuberculosis Care: Single-Arm, Prospective Feasibility Study. JMIR Form Res 2019 Aug 27;3(3):e13411 [FREE Full text] [doi: 10.2196/13411] [Medline: 31456581]
35. Barrientos M, Soria C. Telephones - mobile cellular per capita. IndexMundi. URL: https://www.indexmundi.com/g/r.aspx?v=4010 [accessed 2021-07-19]
36. van Boven JFM, Alffenaar J. An Inca trail to the Holy Grail: digital medication adherence support for TB care. Int J Tuberc Lung Dis 2022 Jan 01;26(1):1-3. [doi: 10.5588/ijtld.21.0613] [Medline: 34969420]

Abbreviations

DHT: digital health technology
DOT: directly observed therapy
EMERGE: ESPACOMP Medication Adherence Reporting Guideline
LTBI: latent tuberculosis infection
M. tb: Mycobacterium tuberculosis
OR: odds ratio
PRISMA: preferred reporting items for systematic review and meta-analysis
RCT: randomized controlled trial
RR: relative risk
TB: tuberculosis
VDOT: video directly observed therapy
VOT: video-observed therapy
WHO: World Health Organization

Edited by T Leung; submitted 26.08.21; peer-reviewed by W Tumuhimbise, R Mpofu, B Vrijens; comments to author 08.10.21; revised version received 30.11.21; accepted 07.01.22; published 23.02.22.

Please cite as:
Ridho A, Alfian SD, van Boven JFM, Levita J, Yalcin EA, Le L, Alffenaar JW, Hak E, Abdulah R, Pradipta IS
Digital Health Technologies to Improve Medication Adherence and Treatment Outcomes in Patients With Tuberculosis: Systematic Review of Randomized Controlled Trials
J Med Internet Res 2022;24(2):e33062
URL: https://www.jmir.org/2022/2/e33062
doi:10.2196/33062
PMID:35195534

Original Paper

The 3-Month Effectiveness of a Stratified Blended Physiotherapy Intervention in Patients With Nonspecific Low Back Pain: Cluster Randomized Controlled Trial

Tjarco Koppenaal[1,2,3], PT, MSc; Martijn F Pisters[1,2,3], PT, PhD; Corelien JJ Kloek[2,4], PT, PhD; Remco M Arensman[2,3], PT, MSc; Raymond WJG Ostelo[5,6], PT, PhD; Cindy Veenhof[2,3,4], PT, PhD

[1]Research Group Empowering Healthy Behaviour, Department of Health Innovations and Technology, Fontys University of Applied Sciences, Eindhoven, Netherlands

[2]Center for Physical Therapy Research and Innovation in Primary Care, Julius Health Care Centers, Utrecht, Netherlands

[3]Physical Therapy Research, Department of Rehabilitation, Physiotherapy Science and Sport, University Medical Center Utrecht Brain Center, Utrecht University, Utrecht, Netherlands

[4]Research Group Innovation of Human Movement Care, Research Center Healthy and Sustainable Living, HU University of Applied Sciences, Utrecht, Netherlands

[5]Department of Health Sciences, Faculty of Science, VU University Amsterdam, Amsterdam Movement Sciences research institute Amsterdam, Amsterdam, Netherlands

[6]Department of Epidemiology and Data Science, Amsterdam University Medical Centre, Location VUmc, Amsterdam, Netherlands

Corresponding Author:
Tjarco Koppenaal, PT, MSc
Research Group Empowering Healthy Behaviour
Department of Health Innovations and Technology
Fontys University of Applied Sciences
Postbus 347, 5600 AH
Eindhoven
Netherlands
Phone: 31 653945086
Email: t.koppenaal@fontys.nl

Abstract

Background: Patient education, home-based exercise therapy, and advice on returning to normal activities are established physiotherapeutic treatment options for patients with nonspecific low back pain (LBP). However, the effectiveness of physiotherapy interventions on health-related outcomes largely depends on patient self-management and adherence to exercise and physical activity recommendations. e-Exercise LBP is a recently developed stratified blended care intervention comprising a smartphone app integrated with face-to-face physiotherapy treatment. Following the promising effects of web-based applications on patients' self-management skills and adherence to exercise and physical activity recommendations, it is hypothesized that e-Exercise LBP will improve patients' physical functioning.

Objective: This study aims to investigate the short-term (3 months) effectiveness of stratified blended physiotherapy (e-Exercise LBP) on physical functioning in comparison with face-to-face physiotherapy in patients with nonspecific LBP.

Methods: The study design was a multicenter cluster randomized controlled trial with intention-to-treat analysis. Patients with nonspecific LBP aged ≥18 years were asked to participate in the study. The patients were treated with either stratified blended physiotherapy or face-to-face physiotherapy. Both interventions were conducted according to the Dutch physiotherapy guidelines for nonspecific LBP. Blended physiotherapy was stratified according to the patients' risk of developing persistent LBP using the Keele STarT Back Screening Tool. The primary outcome was physical functioning (Oswestry Disability Index, range 0-100). Secondary outcomes included pain intensity, fear-avoidance beliefs, and self-reported adherence. Measurements were taken at baseline and at the 3-month follow-up.

Results: Both the stratified blended physiotherapy group (104/208, 50%) and the face-to-face physiotherapy group (104/208, 50%) had improved clinically relevant and statistically significant physical functioning; however, there was no statistically significant or clinically relevant between-group difference (mean difference −1.96, 95% CI −4.47 to 0.55). For the secondary outcomes, stratified blended physiotherapy showed statistically significant between-group differences in fear-avoidance beliefs

and self-reported adherence. In patients with a high risk of developing persistent LBP (13/208, 6.3%), stratified blended physiotherapy showed statistically significant between-group differences in physical functioning (mean difference −16.39, 95% CI −27.98 to −4.79) and several secondary outcomes.

Conclusions: The stratified blended physiotherapy intervention e-Exercise LBP is not more effective than face-to-face physiotherapy in patients with nonspecific LBP in improving physical functioning in the short term. For both stratified blended physiotherapy and face-to-face physiotherapy, within-group improvements were clinically relevant. To be able to decide whether e-Exercise LBP should be implemented in daily physiotherapy practice, future research should focus on the long-term cost-effectiveness and determine which patients benefit most from stratified blended physiotherapy.

Trial Registration: ISRCTN Registry 94074203; https://doi.org/10.1186/ISRCTN94074203

International Registered Report Identifier (IRRID): RR2-https://doi.org/10.1186/s12891-020-3174-z

(J Med Internet Res 2022;24(2):e31675) doi:10.2196/31675

KEYWORDS

eHealth; nonspecific low back pain; physiotherapy; blended care; mobile phone

Introduction

Low back pain (LBP)–related disability and the related socioeconomic burden remain high despite the many treatment options and health care resources available for LBP [1]. LBP can be caused by a specific pathology or trauma; however, in >90% of cases, an underlying disease is absent [2]. The clinical course of this so-called *nonspecific LBP* varies and, as expected, is often less favorable; some patients recover within a couple of days or weeks, and other patients experience persistent disabling symptoms leading to chronic LBP. Up to 65% of primary care patients with LBP still experience pain 1 year after onset [3,4].

Clinical practice guidelines recommend a patient-centered approach for the management of LBP [5,6]. This approach identifies patients with an increased likelihood of delayed recovery at an early stage and stratifies the treatment accordingly [6-8]. An example of a tool for identifying individuals at risk of delayed recovery is the Keele STarT Back Screening Tool [9,10]. In general, in patients who have a *low risk* for delayed recovery, early management comprises advice, reassurance, and education about the nonspecific nature of their LBP and encouragement to stay active. For individuals at *medium risk* for developing persistent LBP, personalized and supervised exercise therapy should be considered. For the *high-risk* group, this exercise therapy can be supported by a graded activity approach or cognitive behavioral components [8,11]. In addition to a patient-centered and stratified approach, patients' adherence to prescribed (home-based) exercises and recommended physical activity behavior is crucial for the effectiveness of care [12]. Earlier research showed that 45% to 70% of patients do not adhere to prescribed exercises and physical activity recommendations, whereas adherent patients with LBP have a reduced risk of recurrent LBP [13,14].

Within the treatment of patients with LBP, *blended care* is a promising new and understudied field [15]. Blended care refers to the integration of web-based and offline components within the treatment process and requires that both components contribute equally to the treatment process [16,17]. The integration of web-based components, such as websites and apps, provides new solutions to monitor and coach patients' individual health behaviors and support the optimization of face-to-face care tailored to the patients' individual needs [18-20]. Thereafter, web-based components can be an effective means of stimulating adherence to prescribed exercises at home between face-to-face sessions and possibly increase self-management of LBP [21,22]. Until now, evidence on patient-centered and stratified care has not been integrated into blended care. Therefore, we recently developed e-Exercise LBP, a stratified blended intervention in which a smartphone app is integrated within face-to-face physiotherapy treatment, and established its feasibility and proof of concept for the treatment of functional disability and pain [23]. e-Exercise LBP is an adapted version of previously developed and evaluated blended physiotherapy programs [24,25]. Following the promising effects of web-based applications for patients' self-management skills and adherence to exercise and physical activity recommendations, it is hypothesized that e-Exercise LBP will improve patients' physical functioning. However, the effectiveness of e-Exercise LBP in comparison with primary care physiotherapy still needs to be determined. The primary aim of this study is to investigate the short-term (3 months) effectiveness of stratified blended physiotherapy (e-Exercise LBP) on physical functioning in comparison with face-to-face physiotherapy in patients with nonspecific LBP.

Methods

Design and Ethical Considerations

The e-Exercise LBP study was a prospective multicenter cluster randomized controlled trial. The study protocol was approved by the medical research ethics committee of the University Medical Center Utrecht, the Netherlands (18-085/D), and registered at the onset of patient enrollment (ISRCTN 94074203). From January 2018 to June 2018, 122 physiotherapists working in 58 primary care physiotherapy practices were recruited and randomized to either stratified blended physiotherapy (e-Exercise LBP) or face-to-face physiotherapy. Details of the design and methods of the study have been published previously [26]. This study is reported according to the CONSORT (Consolidated Standards of Reporting Trials) statement for cluster randomized trials (Multimedia Appendix 1).

Recruitment

Setting and Randomization

Physiotherapists were recruited by an invitational letter sent to the professional network of the authors and physiotherapists who participated in a previous e-Exercise study [24]. In addition, an advertisement was placed in the web-based newsletter of the Royal Dutch Society for Physiotherapy. Physiotherapy practices could participate with ≥1 physiotherapist, regardless of professional experience and education or specialization (eg, manual therapy). Physiotherapists were cluster randomized at the level of practice to avoid contamination. Treatment allocation was concealed and performed by an independent researcher using a computer-generated, a priori created, random sequence table and in a 1:1 allocation ratio. Physiotherapists and patients were not blinded to the group allocation.

The physiotherapists in the stratified blended physiotherapy group received two 4-hour training sessions on e-Exercise LBP and the study procedures. In the face-to-face physiotherapy group, physiotherapists received a 4-hour training session in current best practices according to the LBP guidelines of the Royal Dutch Society for Physiotherapy [11] and the study procedures.

Patients

Patients with LBP who contacted a participating physiotherapy practice were orally informed about the study and invited to participate. Interested patients received a patient information letter by email and an informative phone call by one of the researchers (TK or RMA) before the first appointment. When a patient was willing to participate after the phone call, a face-to-face appointment was scheduled (by TK or RMA) to obtain written informed consent and verify eligibility. The eligibility criteria were as follows: (1) being a patient requesting physiotherapy treatment for nonspecific LBP, defined as pain in the lumbosacral region (sometimes associated with radiating pain to the buttock or leg) [11]; (2) aged ≥18 years; (3) possessing a smartphone or tablet (iOS or Android operating system) with access to the internet; and (4) mastery of the Dutch language. The exclusion criteria were as follows: (1) a specific cause of LBP determined through medical imaging or a medical physician, (2) serious comorbidities (eg, malignancy or stroke), and (3) current pregnancy because of the prevalence of pelvic girdle pain as a specific form of LBP.

Intervention

Experimental: Stratified Blended Physiotherapy (e-Exercise LBP)

Patients allocated to the stratified blended physiotherapy group received blended physiotherapy, comprising a smartphone app integrated within face-to-face physiotherapy treatment [23,26]. Both the contents of the smartphone app and the face-to-face physiotherapy treatment are based on the recommendations of the LBP guidelines of the Royal Dutch Society for Physiotherapy [11]. The duration and content of the stratified blended physiotherapy intervention were based on the patients' risk for developing persistent LBP (*low*, *medium*, or *high*) using the Keele STarT Back Screening Tool [9,10]. The smartphone app contains video-supported self-management information, video-supported exercises, and a goal-oriented physical activity module. Both the contents of face-to-face care and the smartphone app were tailored by the physiotherapists to the patients' individual needs and progress (Table 1). Although physiotherapists were recommended to treat according to the stratified blended physiotherapy protocol, they were free to deviate from the protocol with respect to their clinical competence. Print screens of the smartphone app are provided in Multimedia Appendix 2.

Table 1. Overview of the stratified blended physiotherapy intervention (e-Exercise low back pain [LBP]).

Mode of delivery	Low-risk profile	Medium-risk profile	High-risk profile
Smartphone app			
Duration	3 weeks	12 weeks	12 weeks
Information module	Knowledge-based platform with several LBP self-management information themes (directly available)	12 weekly self-management themes, including assignments	12 weekly self-management themes, including assignments, pain education, and psychosocial risk factors
Exercise module	3 to 4 home-based exercises tailored to the patient's specific functional limitations	3 to 4 home-based exercises tailored to the patient's specific functional limitations	3 to 4 home-based exercises tailored to the patient's specific functional limitations
Physical activity module	Physical activity recommendations in accordance with the LBP guidelines of the Royal Dutch Association for Physiotherapy	A 3-day baseline test to determine the current level of physical activity; an 11-week, 3-times per week, goal-oriented training program to maintain or improve the level of physical activity; in patients avoiding physical activity because of LBP, a graded activity functionality can be activated	A 3-day baseline test to determine the current level of physical activity; an 11-week, 3-times per week, goal-oriented training program to maintain or improve the level of physical activity using a graded activity approach
Face-to-face care			
Sessions	2 sessions	Maximum of 8 sessions	Maximum of 12 sessions
Content	Reassurance, information about LBP, instruction on self-management options, and the importance of adequate physical activity behavior	Content similar to low risk, and in addition, the physiotherapist can consider providing evidence-based interventions (eg, passive or active joint mobilization) as recommended by guideline LBP of the Royal Dutch Association for Physiotherapy	Content similar to medium risk, and in addition, the physiotherapist will address the patient's specific psychosocial risk factors using a cognitive behavioral approach, and pain education will be given
Integration of face-to-face care and smartphone app			
First session	Provide information about LBP and instructions on home-based exercises addressing patient's specific functional limitations using the smartphone app	Provide information about LBP, instructions on home-based exercises addressing patient's specific functional limitations, and instructions on 3-day baseline test using the smartphone app	Provide information about LBP, instructions on home-based exercises addressing patient's specific functional limitations, and instructions on 3-day baseline test using the smartphone app
Middle sessions	N/A[a]	Evaluation of progress with the smartphone app and optimizing face-to-face care	Evaluation of progress with the smartphone app and optimizing face-to-face care
Final session	Evaluate the progress with the smartphone app and give recommendations to prevent recurrent episodes of LBP and maintain or improve the physical activity level	Evaluate the progress with the smartphone app and give recommendations to prevent recurrent episodes of LBP and maintain or improve the physical activity level	Evaluate the progress with the smartphone app and give recommendations to prevent recurrent episodes of LBP and maintain or improve the physical activity level

[a]N/A: not applicable.

Control: Face-to-face Physiotherapy

Patients in the face-to-face physiotherapy group received only face-to-face care following the recommendations of the LBP guidelines of the Royal Dutch Society for Physiotherapy [11]. The guideline distinguishes between three different patient profiles based on the clinical course of recovery (ie, normal recovery, abnormal recovery without predominant psychosocial factors, and abnormal recovery with predominant psychosocial factors) but does not use a specific tool to stratify care a priori. The content of face-to-face physiotherapy was the same as the stratified blended care intervention (ie, information, exercises, and recommendations regarding physical activity). However, no recommendations or restrictions were provided with regard to the number of face-to-face sessions. Although web-based applications, such as websites and apps, are not recommended in the guidelines, physiotherapists were instructed to treat people without using any web-based applications to assure contrast between both groups. Practical content considerations were made by the physiotherapists themselves with respect to their clinical expertise.

Measurements

Patients received a web-based questionnaire and an accelerometer at baseline and after 3 months of follow-up. Baseline measurements were conducted face to face and follow-up measurements through web-based communication (eg, FaceTime) or face to face when requested. No financial incentives were offered to complete the measurements. In the

case of an unfilled questionnaire, patients were reminded after 7 and 14 days.

Outcome Measures

Primary Outcome

Physical functioning because of pain was assessed using the Oswestry Disability Index (ODI; version 2.1a) [27,28]. The ODI was derived from the internationally accepted core outcome set for research into patients with nonspecific LBP [28]. A higher score (0-100) indicates increased functional disability.

Secondary Outcomes

Pain intensity was measured using an 11-point numeric rating scale for the average LBP intensity in the last week (0=no pain and 10=worst possible pain) [28,29].

Physical activity was objectively measured using Activ8 (2M Engineering) [30]. Patients were instructed to wear the Activ8 for 5 consecutive weeks starting at baseline and 8 consecutive days at the 3-month follow-up, except during sleeping, showering, bathing, or swimming. For the purpose of this study, only the first 7 days at both the baseline and 3-month follow-up were used. Accelerometer data were eligible if patients had worn the meter for at least 3 days for ≥10 hours a day [31]. For each patient, the mean time spent in moderate to vigorous physical activity (all activities >3.0 metabolic equivalents [32]) in minutes per day was computed by summation and divided by the number of eligible wearing days.

Fear-avoidance beliefs about physical activity and work were measured using the Fear-Avoidance Beliefs Questionnaire [33]. A higher score (range 0-96) indicates stronger fear and avoidance beliefs about how physical activity and work negatively affect LBP.

Pain catastrophizing was measured using the Pain Catastrophizing Scale [34]. A higher score (range 0-55) indicates a higher level of catastrophizing.

Self-efficacy was measured using the General Self-Efficacy Scale [35,36]. A higher score (range 10-40) indicates greater or stronger perceived self-efficacy.

Self-management ability was assessed using the Dutch version of the short form Patient Activation Measure [37]. A higher score (range 0-100) indicates a higher level of self-management.

Health-related quality of life was measured using the EuroQol-5D-5L [38]. A higher score (range 0-100) indicates a higher health-related quality of life.

Patient self-reported adherence to prescribed home exercises was measured using the Exercise Adherence Rating Scale [39]. A higher score (range 0-24) indicates better adherence.

Other Measures

Physiotherapists were asked to complete a registration form about the number of face-to-face sessions and report the applied treatment modalities per session. Patient characteristics and relevant clinical variables were assessed as part of the baseline questionnaire.

Data Analysis

Overview

Descriptive statistics were used to explore baseline comparability and describe patients' general characteristics, the number of face-to-face physiotherapy sessions, and the treatment modalities. To investigate selective attrition, general characteristics and primary baseline variables of dropouts and nondropouts were compared. All analyses were performed according to the *intention-to-treat* principle. Missing value analyses were performed by assuming the missing at random assumption. Multiple imputation was applied using multivariate imputation by chained equations with predictive mean matching for missing data in all outcomes. A total of 36 imputed data sets were generated, corresponding to the highest missing value percentage [40]. For all analyses, a 2-tailed significance level of $P<.05$ was considered statistically significant.

Analyses of Effectiveness

Linear mixed models (LMMs) with random effects to control for correlation within patients and physiotherapy practices [41] were used to determine the short-term effectiveness of stratified blended physiotherapy compared with face-to-face physiotherapy on primary and secondary outcome measures. Regression coefficients with 95% CIs signifying the differences between stratified blended physiotherapy and face-to-face physiotherapy were estimated. Analyses were adjusted for predefined confounders (eg, age, gender, and duration of pain [42-44]) that changed the between-group estimate by ≥10%. In addition, analyses were also adjusted for variables with a substantial difference at baseline that changed the regression coefficient for the between-group estimate by ≥10%. Potential interaction terms were explored. In the case of a statistically significant interaction term, stratified LMM analyses, controlling for the same variables as the primary analysis, were performed for the effect modifier.

Sample Size

The power calculation was based on the recommendations of Campbell et al [45] for cluster randomized trials and performed for the physical functioning primary outcome at the primary end point of the e-Exercise LBP study (ie, 12-month follow-up). In addition, repeated measures of the primary outcome during follow-up were taken into account [46]. An intraclass correlation coefficient of 0.05 was assumed. In addition, to detect a clinically relevant difference between groups at the 12-month follow-up, a difference of >6 points in physical functioning (ODI) [47,48], and an SD of 14.5 [49] were used in the sample size calculation. For the repeated measures of physical functioning, a correlation of 0.5 was estimated between baseline and follow-up measurements until the 12-month follow-up [46]. On the basis of these assumptions (power 80%; $\alpha=.05$) and an average cluster size of 5, a total of 165 patients were needed. With an expected dropout rate of 20%, a total of 208 participating patients (n=104 per arm) were needed.

Results

Flow of Participants, Therapists, and Centers Through the Study

From June 2018 to December 2019, 434 eligible patients with LBP were asked to participate in 58 physiotherapy practices. In 22 physiotherapy practices allocated to stratified blended physiotherapy and 20 practices allocated to face-to-face physiotherapy, 47.9% (208/434) patients were included (Figure 1).

Baseline characteristics of the patients are presented in Table 2. The stratified blended physiotherapy group comprised more men, more patients with a low level of education, and more patients with a duration of LBP >12 months. No other relevant differences in characteristics were seen between groups. At baseline, complete data on outcome measures were available from 97.1% (101/104) of the patients in the stratified blended physiotherapy group and 99% (103/104) of the patients in the face-to-face physiotherapy group, and eligible accelerometer data were available from 84.6% (88/104) and 83.7% (87/104), respectively. Of the 208 patients, 4 (1.9%) ineligible patients (n=2, 50% in the stratified blended physiotherapy group and 2, 50% in the face-to-face physiotherapy group) were unjustified included, did not receive the allocated intervention and were therefore excluded from all analyses.

At the 3-month follow-up, complete data on outcome measures were available from 86.5% (90/104) of the patients in the stratified blended physiotherapy group and 93.3% (97/104) of the patients in the face-to-face physiotherapy group, and eligible accelerometer data were available from 74% (77/104) and 76% (79/104) of these patients, respectively.

Figure 1. Flow diagram of the e-Exercise low back pain study.

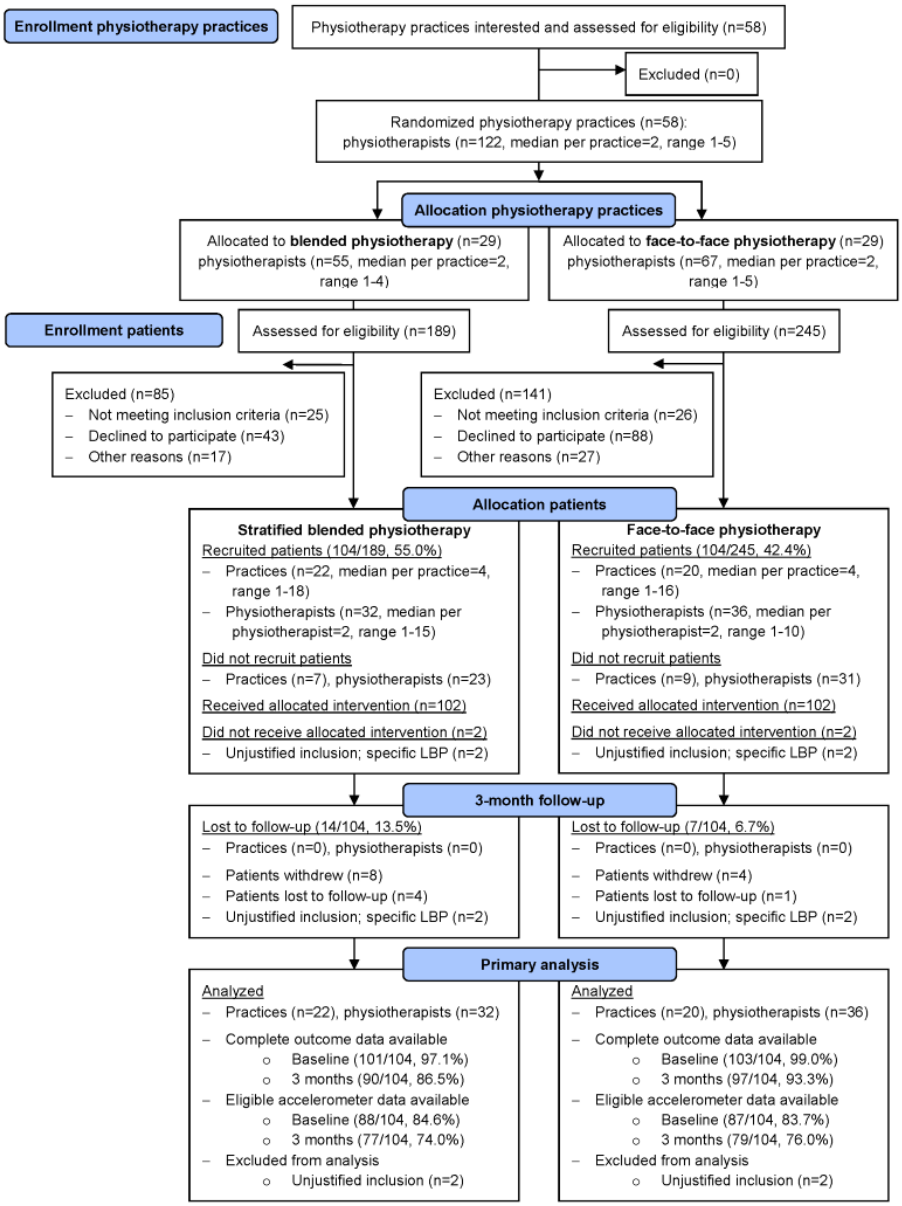


Table 2. Baseline demographic and clinical characteristics for patients from the stratified blended physiotherapy group and face-to-face physiotherapy group (N=208).

Characteristics	Baseline	
	Stratified blended physiotherapy (n=104)	Face-to-face physiotherapy (n=104)
Gender (female), n (%)	45 (43.3)	57 (54.8)
Age (years), mean (SD)	48.10 (15.08)	47.26 (13.58)
BMI (kg/m^2), mean (SD)	25.78 (3.79)	26.31 (5.11)
Presence of comorbidities (yes), n (%)	38 (36.5)	28 (26.9)
Past LBP[a] surgery, n (%)		
None	100 (96.2)	101 (97.1)
Lumbar fusion	0 (0)	1 (1)
Lumbar discectomy	4 (3.9)	2 (1.9)
Central sensitization (score 0-100), mean (SD)	30.88 (13.38)	30.17 (12.19)
Educational level, n (%)		
Low	22 (21.2)	13 (12.5)
Middle	33 (31.7)	36 (34.6)
High	49 (47.1)	55 (52.9)
Duration of LBP complaints, n (%)		
0 to 6 weeks	37 (35.6)	49 (47.1)
6 to 12 weeks	11 (10.6)	19 (18.3)
12 weeks to 12 months	9 (8.7)	9 (8.7)
>12 months	47 (45.2)	27 (26)
Physical functioning (score 0-100), mean (SD)	19.37 (15.64)	20.38 (13.99)
Pain intensity (average score 7 days 0-10), mean (SD)	5.61 (1.99)	5.36 (2.01)
Physical activity (MVPA[b] minutes/day), mean (SD)	80.34 (36.75)	74.82 (40.94)
Health-related quality of life (score 0-100), mean (SD)	67.90 (18.08)	69.75 (17.63)
Fear-avoidance beliefs (score 0-96), mean (SD)	27.86 (16.03)	25.08 (16.18)
Pain catastrophizing (score 0-52), mean (SD)	11.06 (9.30)	10.21 (8.75)
Self-efficacy (score 10-40), mean (SD)	32.13 (4.36)	33.12 (3.62)
Patient activation (score 0-100), mean (SD)	62.48 (12.38)	64.75 (12.68)

[a]LBP: low back pain.

[b]MVPA: moderate to vigorous physical activity.

Number and Treatment Modalities of Physiotherapy Sessions

In total, 189 physiotherapist registration forms were returned (n=95, 50.3% stratified blended physiotherapy and n=94, 49.7% in face-to-face physiotherapy). Table 3 shows the number and treatment modalities of the face-to-face physiotherapy sessions. Patients in the stratified blended physiotherapy group received an average of 4.81 (SD 2.94) face-to-face sessions. For the low-, medium-, and high-risk groups, the average number of sessions was 3.77 (SD 2.54), 5.65 (SD 2.65), and 7.67 (SD 3.54), respectively. Patients in the face-to-face physiotherapy group received an average of 4.94 (SD 2.26) face-to-face sessions. The average number of sessions for the *low-*, *medium-*, and *high-risk* groups was 4.88 (SD 2.02), 5.09 (SD 2.51), and 4.33 (SD 4.16), respectively.

In general, education was the main treatment modality during the face-to-face sessions in both treatment groups. No remarkable differences in treatment modalities were found between the 2 groups or between the different risk groups of developing persistent LBP.

Table 3. Number and treatment modalities of face-to-face physiotherapy sessions for patients from the stratified blended physiotherapy group and face-to-face physiotherapy group.

Category	Stratified blended physiotherapy (risk of developing persistent LBP[a])				Face-to-face physiotherapy (risk of developing persistent LBP)			
	Low (n=52)	Medium (n=34)	High (n=9)	Total (n=95)	Low (n=57)	Medium (n=34)	High (n=3)	Total (n=94)
Number of sessions, mean (SD)	3.77 (2.54)	5.65 (2.65)	7.67 (3.54)	4.81 (2.94)	4.88 (2.02)	5.09 (2.51)	4.33 (4.16)	4.94 (2.26)
Treatment modalities, n (%)[b]								
Education	42 (81)	24 (71)	6 (67)	72 (76)	43 (75)	25 (74)	2 (67)	70 (74)
Strength exercises	9 (17)	3 (9)	1 (11)	13 (14)	7 (12)	6 (18)	0 (0)	13 (14)
Stability exercises	14 (27)	5 (15)	4 (44)	23 (24)	14 (25)	11 (32)	0 (0)	25 (27)
Endurance training	1 (2)	0 (0)	0 (0)	1 (1)	3 (5)	0 (0)	0 (0)	3 (3)
Functional exercises	3 (6)	0 (0)	0 (0)	3 (3)	4 (7)	0 (0)	0 (0)	4 (4)
Active mobilization	15 (29)	10 (29)	2 (22)	27 (28)	22 (39)	11 (32)	2 (67)	35 (37)
Passive mobilization	12 (23)	16 (47)	3 (33)	31 (33)	15 (26)	9 (26)	1 (33)	25 (27)
Massage	4 (8)	8 (24)	2 (22)	14 (15)	9 (19)	5 (15)	0 (0)	14 (15)

[a]LBP: low back pain.

[b]Amount (%) of patients who received the treatment modality as part of the face-to-face physiotherapy session for ≥60% of the total number of face-to-face physiotherapy sessions.

Is Stratified Blended Physiotherapy Effective Compared With Face-to-face Physiotherapy?

In the mixed model analyses, log likelihood ratios of naive models and models that included a random intercept for both physiotherapy practice and physiotherapist were similar. Therefore, physiotherapy practice or physiotherapist was not included as a level in the LMM analyses. At 3 months, LMM analyses showed no clinically relevant or statistically significant between-group difference in the primary outcome of physical functioning (mean difference [MD] −1.96, 95% CI −4.47 to 0.55). For the secondary outcomes, a statistically significant between-group difference was found in favor of stratified blended physiotherapy for fear-avoidance beliefs (MD −4.29, 95% CI −7.22 to −1.37) and patients' self-reported adherence to prescribed home exercises (MD 0.73, 95% CI 0.06-1.39). Within-group analyses showed clinically relevant and statistically significant improvements in physical functioning (MD −11.48, 95% CI −15.06 to −7.91), average pain intensity (MD −2.38, 95% CI −3.00 to −1.76), and fear-avoidance beliefs (MD −5.14, 95% CI −9.22 to −1.06) in the stratified blended physiotherapy group. In the face-to-face physiotherapy group, clinically relevant and statistically significant improvements in physical functioning (MD −11.22, 95% CI −14.64 to −7.80) and average pain intensity (MD −2.51, 95% CI −3.11 to −1.90) were found (Table 4).

As indicated by a statistically significant interaction term, the patients' risk of developing persistent LBP was an effect modifier of the between-group differences on the primary outcome of physical functioning. In patients with a high risk of developing persistent LBP, the stratified analysis showed a statistically significant between-group difference in favor of stratified blended physiotherapy on physical functioning (MD –16.39, 95% CI –27.98 to –4.79), average pain intensity (MD –3.43, 95% CI –6.55 to –0.31), and fear-avoidance beliefs (MD –14.51, 95% CI –28.21 to –0.81). In patients with a medium risk of developing persistent LBP, a statistically significant between-group difference was found in favor of stratified blended physiotherapy on fear-avoidance beliefs (MD –5.93, 95% CI –11.45 to –0.40). In patients with a low risk of developing persistent LBP, no statistically significant between-group differences were found (Table 5).

Table 4. Unadjusted and adjusted primary and secondary outcome measures: improvements and differences within and between groups (N=204).

Stratified blended physiotherapy (n=102)				Face-to-face physiotherapy (n=102)				Between group differences[a]			
Measurements, mean (SD)		Unadjusted within-group differences		Measurements, mean (SD)		Unadjusted within-group differences		Unadjusted		Adjusted[b]	
Baseline	3 months	Mean (95% CI)	*P* value	Baseline	3 months	Mean (95% CI)	*P* value	Mean (95% CI)	*P* value	Mean (95% CI)	*P* value
Physical functioning (range 0-100)											
19.39 (15.56)	7.91 (9.64)	−11.48 (−15.06 to −7.91)	<.001	20.20 (13.90)	8.97 (10.75)	−11.22 (−14.64 to −7.80)	<.001	−0.83 (−3.43 to 1.77)	.53	−1.98 (−4.49 to 0.53)	.12
Pain intensity (average score 7 days; range 0-10)											
5.67 (1.94)	3.29 (2.42)	−2.38 (−3.00 to −1.75)	<.001	5.40 (2.00)	2.90 (2.36)	−2.51 (−3.11 to −1.90)	<.001	0.31 (−0.35 to 0.98)	.36	0.08 (−0.57 to 0.74)	.80
Physical activity (MVPA[c] min/day)											
81.97 (38.52)	78.58 (44.45)	−3.37 (−15.63 to 8.88)	.59	75.70 (41.89)	71.24 (40.34)	−4.42 (−16.91 to 8.07)	.49	3.49 (−8.38 to 15.36)	.56	3.62 (−8.27 to 15.51)	.55
Fear-avoidance beliefs (range 0-96)											
27.92 (16.01)	22.77 (13.38)	−5.14 (−9.25 to −1.04)	.01	25.51 (16.24)	24.82 (16.92)	−0.70 (−5.26 to 3.87)	.77	−3.73 (−6.63 to −0.82)	.01	−4.29 (−7.22 to −1.37)	<.001
Pain catastrophizing (range 0-52)											
11.02 (9.30)	8.97 (8.05)	−2.04 (−4.50 to 0.43)	.11	10.33 (8.76)	9.16 (9.84)	−1.17 (−3.74 to 1.40)	.37	−0.63 (−2.58 to 1.32)	.53	−0.96 (−2.95 to 1.02)	.34
Self-efficacy (range 10-40)											
32.05 (4.38)	32.02 (4.27)	−0.03 (−1.24 to 1.19)	.97	33.12 (3.63)	32.58 (3.99)	−0.54 (−1.59 to 0.52)	.32	0.12 (−0.82 to 1.06)	.81	0.14 (−0.82 to 1.10)	.77
Health-related quality of life (range 0-100)											
67.70 (18.09)	71.44 (20.07)	3.73 (−1.68 to 9.14)	.18	69.75 (17.71)	72.57 (21.06)	2.82 (−2.56 to 8.20)	.31	−0.65 (−6.38 to 5.08)	.82	0.95 (−4.80 to 6.69)	.75
Patient activation (range 0-100)											
62.43 (12.37)	62.45 (11.89)	0.02 (−3.42 to 3.46)	.99	64.72 (12.65)	64.39 (12.71)	−0.33 (−3.84 to 3.18)	.85	−0.83 (−3.94 to 2.27)	.60	−0.79 (−3.95 to 2.36)	.62
Adherence to prescribed home exercises (range 0-24)[d]											
N/A[e]	11.96 (2.43)	N/A	N/A	N/A	11.18 (2.17)	N/A	N/A	0.78 (0.13 to 1.44)	.02	0.73 (0.06 to 1.39)	.03

[a]Difference between baseline and 3 months in stratified blended physiotherapy versus face-to-face physiotherapy.

[b]Adjusted for baseline and duration of low back pain complaints (<12 vs >12 weeks).

[c]MVPA: moderate to vigorous physical activity.

[d]Patient self-reported adherence to prescribed home exercises could only be measured after the treatment period.

[e]N/A: not applicable.

Table 5. Adjusted primary and secondary outcome measures: improvements and differences between groups stratified for the risk of developing persistent low back pain (LBP; N=204).

Outcome measure	Risk of developing persistent LBP					
	Low risk (n=120)		Medium risk (n=71)		High risk (n=13)	
	Between-group difference, mean (95% CI)[a]	*P* value	Between-group difference, mean (95% CI)[a]	*P* value	Between-group difference, mean (95% CI)[a]	*P* value
Physical functioning (range 0-100)	−0.82 (−2.92 to 1.27)	.44	−3.48 (−8.99 to 2.03)	.22	−16.39 (−27.98 to −4.79)	.01
Pain intensity (average score 7 days; range 0-10)	0.30 (−0.52 to 1.13)	.47	0.01 (−1.08 to 1.11)	.98	−3.43 (−6.55 to −0.31)	.03
Physical activity (MVPA[b] minutes/day)	3.80 (−12.05 to 19.65)	.64	1.08 (−16.70 to 18.86)	.91	39.50 (−1.24 to 80.24)	.06
Fear-avoidance beliefs (range 0-96)	−2.70 (−6.22 to 0.82)	.13	−5.93 (−11.45 to −0.40)	.04	−14.51 (−28.21 to −0.81)	.04
Pain catastrophizing (range 0-52)	0.28 (−2.03 to 2.59)	.81	−2.66 (−5.73 to 0.41)	.09	−14.47 (−31.89 to 2.94)	.10
Self-efficacy (range 10-40)	−0.58 (−1.76 to 0.60)	.33	0.85 (−0.92 to 2.62)	.35	1.50 (−4.02 to 7.02)	.60
Health-related quality of life (range 0-100)	1.26 (−7.15 to 9.68)	.77	0.84 (−6.47 to 8.15)	.82	15.84 (−3.92 to 35.61)	.12
Patient activation (range 0-100)	−2.22 (−6.38 to 1.93)	.29	1.85 (−3.27 to 6.97)	.48	7.49 (−1.35 to 16.34)	.10
Adherence to prescribed home exercises (range 0-24)	0.82 (−0.01 to 1.65)	.05	0.86 (−0.35 to 2.08)	.16	−1.19 (−3.37 to 0.99)	.28

[a]Difference between baseline and 3 months in stratified blended physiotherapy versus face-to-face physiotherapy per risk group and adjusted for baseline and duration of low back pain complaints (<12 vs >12 weeks).

[b]MVPA: moderate to vigorous physical activity.

Discussion

Principal Findings

This study evaluated the short-term (3 months) effectiveness of the stratified blended physiotherapy intervention e-Exercise LBP on physical functioning in comparison with face-to-face physiotherapy in patients with nonspecific LBP. In contrast to our expectations, the study results showed no statistically significant between-group difference in physical functioning and most of the secondary outcome measures. Only fear-avoidance beliefs and patient self-reported adherence to prescribed home exercises improved significantly in patients who were allocated to stratified blended physiotherapy. When looking at the different prognostic risk groups in patients with a high risk of developing persistent LBP, a statistically significant between-group difference in favor of stratified blended physiotherapy on physical functioning, average pain intensity, and fear-avoidance beliefs was found; however, these results come with some uncertainty.

Interpretation of the Findings

The results of this study complement the findings from previous systematic reviews of randomized controlled trials that showed that in the short term, web-based applications could reduce LBP-related pain and disability; however, when compared with other interventions, the results are inconclusive [15,22,50]. A possible explanation for these inconclusive findings is the considerable heterogeneity in the studied characteristics and comparators, which hampers a clear comparison. For example, in our study, we integrated a web-based application within face-to-face guidance and compared it with face-to-face physiotherapy. Previous studies in this research area have focused predominantly on web-based applications as a stand-alone intervention without the face-to-face guidance of a health care professional [15,22,50]. Only a few studies have investigated web-based applications as an adjunct to face-to-face guidance, and the results regarding the added value of these combined interventions have been inconclusive [15,51]. Similar to our study, Sandal et al [51] investigated a smartphone app as an adjunct to face-to-face guidance. The app was tailored using artificial intelligence and did not influence face-to-face guidance. In this study, the reported between-group difference was statistically significant in favor of the combined intervention when compared with face-to-face guidance alone; however, the difference was small and of uncertain clinical significance.

Another example of heterogeneity in research on web-based applications is the large variation in delivery modes and duration. Similar to e-Exercise LBP, most web-based applications tailored the content of the intervention using patient characteristics and focused on self-management support,

home-based exercise, and physical activity prescription [15,22,50]. However, the e-Exercise LBP app provided this content in weekly information modules and daily reminders to exercise and physical activity recommendations during a 3- or 12-week duration [26]; the duration in other studies ranged from 3 weeks to 1 year. In addition, the delivery modes showed large variation; that is, from no specific recommendations to multiple web- or telephone-based coaching sessions [15,22,50].

Thus, looking at the different characteristics of web-based applications, such as the role of the health care professional within the intervention and the delivery mode and duration, future research needs to focus on the comparison of web-based applications with different characteristics to obtain a better understanding of which elements work the best.

In our study, the short-term within-group improvements in physical functioning and average pain intensity of stratified blended physiotherapy were comparable with face-to-face physiotherapy, both of which were statistically significant and clinically meaningful. Patients in the stratified blended physiotherapy group improved on average 11.48 (95% CI −15.06 to −7.91) points (59.5%) in physical functioning, and patients in the face-to-face physiotherapy group improved by an average of 11.22 (95% CI −14.64 to −7.80) points (56%). For average pain intensity, these improvements were 2.38 (95% CI −3.00 to −1.76) points (42.8%) and 2.51 (95% CI −3.11 to −1.90) points (46.9%), respectively. As physical functioning and average pain intensity decreased by >30%, the improvements in both groups were considered clinically meaningful [52]. At the moment, e-Exercise LBP cannot be considered an alternative to face-to-face physiotherapy as this study was conducted as a superiority trial. To be able to value the true potential of e-Exercise LBP, the meaningful within-group improvements must be considered from the perspective of the additional effort and costs needed to implement such an intervention in daily physiotherapy practice. Future cost-effectiveness analyses will provide more insight into the long-term economic benefits of stratified blended physiotherapy. On the other hand, given the additional effort and costs, the potential of e-Exercise LBP needs to be considered from the perspective of future health care. It is expected that technology will be increasingly integrated into care for patients who are suitable to use it. Future studies need to determine which patients benefit most from a stratified blended physiotherapy approach.

The e-Exercise LBP intervention significantly increased patients' self-reported adherence to prescribed home exercises, as hypothesized. In addition, it resulted in a significant reduction of fear-avoidance beliefs when compared with face-to-face physiotherapy. The between-group difference in patients' self-reported adherence to prescribed home exercises was 3.3% points in favor of the e-Exercise LBP intervention. For fear-avoidance beliefs, the between-group difference was −4.6% points in favor of the e-Exercise LBP intervention. Although there are no established cutoffs for the minimum clinically important between-group differences in these outcomes, we consider the between-group differences as small. The difference in adherence might be explained by the benefits of integrating a smartphone app. The 24/7 availability of the app and functionality to remind the patient to perform scheduled exercises might have stimulated the patients to adhere to their prescribed home exercises in a better way than in the face-to-face physiotherapy group [18,53]. Further research on the long-term clinical relevance of adherence to home exercises as prescribed in e-Exercise LBP is ongoing.

The reduction of fear-avoidance beliefs complements evidence from a systematic review and meta-analysis that concluded that patient education provides reassurance for patients with acute or subacute LBP [54]. In our study, this reduction in the stratified blended physiotherapy group might be explained by the information module of the smartphone app. As the information module provides the patient with self-management information about LBP, the patient can reread the advice and reassurance given in the face-to-face sessions by the physiotherapist about their LBP at all times. As a result, the harmless and nonspecific nature of LBP is possibly remembered in a better way [55]. Long-term results should indicate whether this reduction in fear-avoidance beliefs also influences physical functioning, the handling of recurrent complaints, and costs a patient incurs because of LBP.

Several explanations are possible to clarify why the additional benefits of stratified blended physiotherapy were not found. A first explanation is that the added value of a stratified approach in itself must be critically evaluated. Although clinical practice guidelines have adopted and advocated a stratified care approach for several years to improve patient outcomes, the added value of this approach is, at present, unclear. On the basis of previous recommendations, we decided to use the Keele STarT Back Screening Tool to create a matched web-based application [10]. Our results show that, after specific training, treatment intensity (ie, the number of face-to-face sessions) in the e-Exercise LBP group was in line with the patient's risk profile, which was not the case in our control group. However, this difference in treatment intensity did not lead to relevant between-group differences. This seems to be in line with more recent studies evaluating the stratified approach according to the Keele STarT Back Screening Tool. The results from these studies are not convincing regarding the added value of such a stratified approach [56,57]. Future research should focus on determining whether this concerns the added value of the tool itself or the added value of a stratified care approach in general.

In addition, stratified blended physiotherapy might not be suitable for every patient. Earlier research has shown that it is difficult to determine what works best for each individual patient [22,50]. In our study, we did not take into account the patient's suitability for blended care to determine the optimal personalized blended treatment [58]. As a result, patients might have received stratified blended physiotherapy without being suitable for it; for example, a lack of motivation or digital literacy skills. Consequently, this could have resulted in the suboptimal effectiveness of our stratified blended physiotherapy intervention when compared with face-to-face physiotherapy. For future studies on blended care, it is recommended to use patients' suitability for blended care as inclusion criteria or criteria to match treatment. The Dutch Blended Physiotherapy Checklist [58] could be a useful aid in this process.

A third explanation might be the relatively high proportion of patients with a low risk of developing persistent LBP in this study. For this group, earlier research has shown that providing advice as a single intervention is likely to reassure the patient with LBP but does not result in different management of pain and disability in the short term [54,59]. In addition, for this group, a stratified approach is beneficial from an economic perspective rather than in terms of clinical outcomes, as many of these patients recover completely within 2 to 3 weeks but nevertheless receive unnecessary treatment [57,60,61].

A final explanation is the timing of our follow-up measurement at 3 months only. Given the favorable course of LBP [62] and the rationale that stratified blended physiotherapy will stimulate patients' self-management and adherence [21,22], patients in the stratified blended physiotherapy group might recover faster, which is not captured by a single follow-up measurement at 3 months. Therefore, for future studies that aim to investigate postintervention effectiveness, it is recommended to measure the clinical outcomes immediately after the intervention is completed and to monitor the time to recovery.

Strengths and Limitations

This study had several important strengths. It is the next step in a multiphase development and implementation process based on the Center for eHealth Research Roadmap [63]. After developing a prototype and testing its feasibility in a pilot study [23], this study determined the short-term effectiveness of the final stratified blended physiotherapy protocol and showed its potential compared with face-to-face physiotherapy. The pragmatic, multicenter, cluster randomized controlled trial design allowed for the evaluation of stratified, blended physiotherapy in comparison with face-to-face physiotherapy in a real-world situation. The baseline characteristics of both treatment groups and the distribution of the different prognostic risk groups of developing persistent LBP reflect the characteristics of patients with LBP normally being treated in primary care physiotherapy [60], which enhances the generalizability of our results. The use of measurement instruments recommended in the core outcome set for research into patients with nonspecific LBP [28] and a low dropout rate (10.1%) guaranteed the internal validity of the results.

Nevertheless, this study also had a few limitations. First, the results seem to suggest that patients' risk of developing persistent LBP could be an effect modifier of the between-group differences on the primary outcome. Especially in the highest risk group, consistent between-group differences were seen in both the primary and secondary outcomes, supporting the rationale for stratified blended physiotherapy. As it was not the primary aim of this study, the sample size calculation did not take interaction into account, the numbers were small, and therefore, the results should be interpreted with caution. Second, as we conducted a pragmatic study, the experiences of physiotherapists in either using web-based applications or treating patients with nonspecific LBP were not considered inclusion criteria for physiotherapy practices. However, given both the complexity of blended care [17] and the complexity of treating patients with nonspecific LBP [4], it can be expected that more experienced physiotherapists are able to deliver better treatment than less experienced physiotherapists. Therefore, experience might have influenced our analysis. Finally, 4 included patients were excluded from the analysis after being diagnosed with specific LBP. As this number is low and occurred equally in both treatment groups (2 in each group), we expect that this has not influenced the results [64].

Conclusions

The stratified blended physiotherapy intervention e-Exercise LBP is not more effective than face-to-face physiotherapy in patients with nonspecific LBP in improving physical functioning in the short term. For both stratified blended physiotherapy and face-to-face physiotherapy, within-group improvements were clinically relevant. To be able to decide whether e-Exercise LBP should be implemented in daily physiotherapy practice, future research should focus on the long-term cost-effectiveness and determine which patients benefit most from stratified blended physiotherapy.

Acknowledgments

This study was cofunded by the Taskforce for Applied Research SIA (RAAK.PRO02.063), part of the Dutch Research Council. The funder had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors are grateful for the invaluable help provided by Loes Kolenberg in the analysis of the physical activity data.

Authors' Contributions

TK participated in the development of the design of the study, conducted the study, performed the statistical analyses, and drafted and revised the manuscript. MFP was the principal investigator, developed the design of the study, and participated in the writing and revision of the manuscript. CJJK participated in the development of the design of the study and participated in the writing and revision of the manuscript. RMA participated in the design of the study, conducted the study, and participated in revising the manuscript. RWJGO supervised the project, advised the design of the study and statistical analysis, and participated in revising the manuscript. CV supervised the project, advised the design of the study and statistical analysis, and participated in revising the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth; version 1.6) checklist.
[PDF File (Adobe PDF File), 2486 KB - jmir_v24i2e31675_app1.pdf]

Multimedia Appendix 2
Print screens of the smartphone app.
[PDF File (Adobe PDF File), 307 KB - jmir_v24i2e31675_app2.pdf]

References

1. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1789-1858 [FREE Full text] [doi: 10.1016/S0140-6736(18)32279-7] [Medline: 30496104]
2. Koes BW, van Tulder MW, Thomas S. Diagnosis and treatment of low back pain. BMJ 2006 Jun 17;332(7555):1430-1434 [FREE Full text] [doi: 10.1136/bmj.332.7555.1430] [Medline: 16777886]
3. Itz C, Geurts J, van Kleef M, Nelemans P. Clinical course of non-specific low back pain: a systematic review of prospective cohort studies set in primary care. Eur J Pain 2013 Jan;17(1):5-15. [doi: 10.1002/j.1532-2149.2012.00170.x] [Medline: 22641374]
4. Hartvigsen J, Hancock MJ, Kongsted A, Louw Q, Ferreira ML, Genevay S, Lancet Low Back Pain Series Working Group. What low back pain is and why we need to pay attention. Lancet 2018 Jun 09;391(10137):2356-2367. [doi: 10.1016/S0140-6736(18)30480-X] [Medline: 29573870]
5. Lin I, Wiles L, Waller R, Goucke R, Nagree Y, Gibberd M, et al. What does best practice care for musculoskeletal pain look like? Eleven consistent recommendations from high-quality clinical practice guidelines: systematic review. Br J Sports Med 2020 Jan;54(2):79-86. [doi: 10.1136/bjsports-2018-099878] [Medline: 30826805]
6. Foster NE, Anema JR, Cherkin D, Chou R, Cohen SP, Gross DP, et al. Prevention and treatment of low back pain: evidence, challenges, and promising directions. The Lancet 2018 Jun;391(10137):2368-2383. [doi: 10.1016/s0140-6736(18)30489-6]
7. Maher C, Underwood M, Buchbinder R. Non-specific low back pain. Lancet 2017 Feb 18;389(10070):736-747. [doi: 10.1016/S0140-6736(16)30970-9] [Medline: 27745712]
8. Wong J, Côté P, Sutton D, Randhawa K, Yu H, Varatharajan S, et al. Clinical practice guidelines for the noninvasive management of low back pain: a systematic review by the Ontario Protocol for Traffic Injury Management (OPTIMa) Collaboration. Eur J Pain 2017 Feb;21(2):201-216. [doi: 10.1002/ejp.931] [Medline: 27712027]
9. Hill JC, Whitehurst DG, Lewis M, Bryan S, Dunn KM, Foster NE, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. Lancet 2011 Oct;378(9802):1560-1571. [doi: 10.1016/S0140-6736(11)60937-9]
10. Hill JC, Dunn KM, Lewis M, Mullis R, Main CJ, Foster NE, et al. A primary care back pain screening tool: identifying patient subgroups for initial treatment. Arthritis Rheum 2008 May 15;59(5):632-641 [FREE Full text] [doi: 10.1002/art.23563] [Medline: 18438893]
11. Staal JB, Hendriks EJ, Heijmans M, Kiers H, Rutten AM, Tulder MW. KNGF-richtlijn Lage rugpijn. KNGF-richtlijn Lage rugpijn 2013;7.
12. Friedrich M, Gittler G, Halberstadt Y, Cermak T, Heiller I. Combined exercise and motivation program: effect on the compliance and level of disability of patients with chronic low back pain: a randomized controlled trial. Archives Physical Med Rehab 1998 May;79(5):475-487. [doi: 10.1016/s0003-9993(98)90059-4]
13. van Gool CH, Penninx BW, Kempen GI, Rejeski WJ, Miller GD, van Eijk JT, et al. Effects of exercise adherence on physical function among overweight older adults with knee osteoarthritis. Arthritis Rheum 2005 Feb 15;53(1):24-32 [FREE Full text] [doi: 10.1002/art.20902] [Medline: 15696558]
14. Mannion AF, Helbling D, Pulkovski N, Sprott H. Spinal segmental stabilisation exercises for chronic low back pain: programme adherence and its influence on clinical outcome. Eur Spine J 2009 Dec;18(12):1881-1891 [FREE Full text] [doi: 10.1007/s00586-009-1093-7] [Medline: 19609785]
15. Dario AB, Moreti Cabral A, Almeida L, Ferreira ML, Refshauge K, Simic M, et al. Effectiveness of telehealth-based interventions in the management of non-specific low back pain: a systematic review with meta-analysis. Spine J 2017 Sep;17(9):1342-1351. [doi: 10.1016/j.spinee.2017.04.008] [Medline: 28412562]
16. van der Vaart R, Witting M, Riper H, Kooistra L, Bohlmeijer ET, van Gemert-Pijnen LJ. Blending online therapy into regular face-to-face therapy for depression: content, ratio and preconditions according to patients and therapists using a Delphi study. BMC Psychiatry 2014 Dec 14;14:355 [FREE Full text] [doi: 10.1186/s12888-014-0355-z] [Medline: 25496393]
17. Wentzel J, van der Vaart R, Bohlmeijer ET, van Gemert-Pijnen JE. Mixing online and face-to-face therapy: how to benefit from blended care in mental health care. JMIR Ment Health 2016 Feb 09;3(1):e9 [FREE Full text] [doi: 10.2196/mental.4534] [Medline: 26860537]

18. Lustria ML, Cortese J, Noar SM, Glueckauf RL. Computer-tailored health interventions delivered over the web: review and analysis of key components. Patient Educ Couns 2009 Feb;74(2):156-173. [doi: 10.1016/j.pec.2008.08.023] [Medline: 18947966]
19. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. Psychol Bull 2007 Jul;133(4):673-693. [doi: 10.1037/0033-2909.133.4.673] [Medline: 17592961]
20. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of web-based vs. non-web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40 [FREE Full text] [doi: 10.2196/jmir.6.4.e40] [Medline: 15631964]
21. Bennell KL, Marshall CJ, Dobson F, Kasza J, Lonsdale C, Hinman RS. Does a web-based exercise programming system improve home exercise adherence for people with musculoskeletal conditions?: a randomized controlled trial. Am J Phys Med Rehabil 2019 Oct;98(10):850-858. [doi: 10.1097/PHM.0000000000001204] [Medline: 31021823]
22. Nicholl BI, Sandal LF, Stochkendahl MJ, McCallum M, Suresh N, Vasseljen O, et al. Digital support interventions for the self-management of low back pain: a systematic review. J Med Internet Res 2017 May 21;19(5):e179 [FREE Full text] [doi: 10.2196/jmir.7290] [Medline: 28550009]
23. Kloek C, van Tilburg M, Staal J, Veenhof C, Bossen D. Development and proof of concept of a blended physiotherapeutic intervention for patients with non-specific low back pain. Physiotherapy 2019 Dec;105(4):483-491 [FREE Full text] [doi: 10.1016/j.physio.2018.12.006] [Medline: 31031023]
24. Kloek CJ, Bossen D, Spreeuwenberg PM, Dekker J, de Bakker DH, Veenhof C. Effectiveness of a blended physical therapist intervention in people with hip osteoarthritis, knee osteoarthritis, or both: a cluster-randomized controlled trial. Phys Ther 2018 Jul 01;98(7):560-570 [FREE Full text] [doi: 10.1093/ptj/pzy045] [Medline: 29788253]
25. Kloek CJ, van Dongen JM, de Bakker DH, Bossen D, Dekker J, Veenhof C. Cost-effectiveness of a blended physiotherapy intervention compared to usual physiotherapy in patients with hip and/or knee osteoarthritis: a cluster randomized controlled trial. BMC Public Health 2018 Aug 31;18(1):1082 [FREE Full text] [doi: 10.1186/s12889-018-5975-7] [Medline: 30170586]
26. Koppenaal T, Arensman RM, van Dongen JM, Ostelo RW, Veenhof C, Kloek CJ, et al. Effectiveness and cost-effectiveness of stratified blended physiotherapy in patients with non-specific low back pain: study protocol of a cluster randomized controlled trial. BMC Musculoskelet Disord 2020 Apr 22;21(1):265 [FREE Full text] [doi: 10.1186/s12891-020-3174-z] [Medline: 32321492]
27. Fairbank JC, Pynsent PB. The Oswestry disability index. Spine (Phila Pa 1976) 2000 Nov 15;25(22):2940-52; discussion 2952. [doi: 10.1097/00007632-200011150-00017] [Medline: 11074683]
28. Chiarotto A, Terwee CB, Ostelo RW. Choosing the right outcome measurement instruments for patients with low back pain. Best Pract Res Clin Rheumatol 2016 Dec;30(6):1003-1020. [doi: 10.1016/j.berh.2017.07.001] [Medline: 29103546]
29. Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain 2008 Feb;9(2):105-121. [doi: 10.1016/j.jpain.2007.09.005] [Medline: 18055266]
30. Horemans H, Kooijmans H, van den Berg-Emons R, Bussmann H. The Activ8 activity monitor: validation of posture and movement classification. J Rehabil Assist Technol Eng 2020;7:2055668319890535 [FREE Full text] [doi: 10.1177/2055668319890535] [Medline: 32206335]
31. Quante M, Kaplan ER, Rueschman M, Cailler M, Buxton OM, Redline S. Practical considerations in using accelerometers to assess physical activity, sedentary behavior, and sleep. Sleep Health 2015 Dec;1(4):275-284. [doi: 10.1016/j.sleh.2015.09.002] [Medline: 29073403]
32. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Tudor-Locke C, et al. 2011 Compendium of Physical Activities: a second update of codes and MET values. Med Sci Sports Exerc 2011 Aug;43(8):1575-1581. [doi: 10.1249/MSS.0b013e31821ece12] [Medline: 21681120]
33. Waddell G, Newton M, Henderson I, Somerville D, Main C. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. Pain 1993 Feb;52(2):157-168. [doi: 10.1016/0304-3959(93)90127-B] [Medline: 8455963]
34. Sullivan MJ, Bishop SR, Pivik J. The pain catastrophizing scale: development and validation. Psychol Assessment 1995;7(4):524-532. [doi: 10.1037/1040-3590.7.4.524]
35. Teeuw B, Schwarzer R, Jerusalem M. Dutch adaptation of the general self-efficacy scale. Berlin. 1994. URL: http://userpage.fu-berlin.de/~health/dutch.htm [accessed 2022-02-10]
36. Scholz U, Dona B, Sud S, Schwarzer R. Is general self- efficacy a universal construct? Psychometric findings from 25 countries. Eur J Psychol Assess 2002;18(3):42-51. [doi: 10.1027/1015-5759.18.3.242]
37. Rademakers J, Nijman J, van der Hoek L, Heijmans M, Rijken M. Measuring patient activation in The Netherlands: translation and validation of the American short form Patient Activation Measure (PAM13). BMC Public Health 2012 Jul 31;12:577 [FREE Full text] [doi: 10.1186/1471-2458-12-577] [Medline: 22849664]
38. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]

39. Newman-Beinart NA, Norton S, Dowling D, Gavriloff D, Vari C, Weinman JA, et al. The development and initial psychometric evaluation of a measure assessing adherence to prescribed exercise: the Exercise Adherence Rating Scale (EARS). Physiotherapy 2017 Jun;103(2):180-185. [doi: 10.1016/j.physio.2016.11.001] [Medline: 27913064]
40. Sullivan TR, Salter AB, Ryan P, Lee KJ. Bias and precision of the "Multiple imputation, then deletion" method for dealing with missing outcome data. Am J Epidemiol 2015 Sep 15;182(6):528-534. [doi: 10.1093/aje/kwv100] [Medline: 26337075]
41. Goldstein H. Multilevel Statistical Models. Hoboken, New Jersey, United States: Wiley; 1987.
42. Dunn KM, Jordan K, Croft P. Contributions of prognostic factors for poor outcome in primary care low back pain patients. Eur J Pain 2011 Mar;15(3):313-319 [FREE Full text] [doi: 10.1016/j.ejpain.2010.07.008] [Medline: 20728385]
43. Hayden J, Chou R, Hogg-Johnson S, Bombardier C. Systematic reviews of low back pain prognosis had variable methods and results: guidance for future prognosis reviews. J Clin Epidemiol 2009 Aug;62(8):781-96.e1. [doi: 10.1016/j.jclinepi.2008.09.004] [Medline: 19136234]
44. Grotle M, Foster N, Dunn K, Croft P. Are prognostic indicators for poor outcome different for acute and chronic low back pain consulters in primary care? Pain 2010 Dec;151(3):790-797 [FREE Full text] [doi: 10.1016/j.pain.2010.09.014] [Medline: 20932646]
45. Campbell MK, Piaggio G, Elbourne DR, Altman DG, CONSORT Group. Consort 2010 statement: extension to cluster randomised trials. BMJ 2012 Sep 04;345:e5661. [doi: 10.1136/bmj.e5661] [Medline: 22951546]
46. Vickers AJ. How many repeated measures in repeated measures designs? Statistical issues for comparative trials. BMC Med Res Methodol 2003 Oct 27;3:22 [FREE Full text] [doi: 10.1186/1471-2288-3-22] [Medline: 14580266]
47. Fritz JM, Irrgang J. A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. Phys Ther 2001 Feb;81(2):776-788. [doi: 10.1093/ptj/81.2.776] [Medline: 11175676]
48. Resnik L, Dobrzykowski E. Guide to outcomes measurement for patients with low back pain syndromes. J Orthop Sports Phys Ther 2003 Jun;33(6):307-16; discussion 317. [doi: 10.2519/jospt.2003.33.6.307] [Medline: 12839205]
49. Apeldoorn AT, Ostelo RW, van Helvoirt H, Fritz JM, Knol DL, van Tulder MW, et al. A randomized controlled trial on the effectiveness of a classification-based system for subacute and chronic low back pain. Spine 2012;37(16):1347-1356. [doi: 10.1097/brs.0b013e31824d9f2b]
50. Du S, Liu W, Cai S, Hu Y, Dong J. The efficacy of e-health in the self-management of chronic low back pain: a meta analysis. Int J Nurs Stud 2020 Jun;106:103507. [doi: 10.1016/j.ijnurstu.2019.103507] [Medline: 32320936]
51. Sandal LF, Bach K, Øverås CK, Svendsen MJ, Dalager T, Stejnicher Drongstrup Jensen J, et al. Effectiveness of app-delivered, tailored self-management support for adults with lower back pain-related disability: a selfBACK randomized clinical trial. JAMA Intern Med 2021 Oct 01;181(10):1288-1296 [FREE Full text] [doi: 10.1001/jamainternmed.2021.4097] [Medline: 34338710]
52. Ostelo RW, Deyo RA, Stratford P, Waddell G, Croft P, Von Korff M, et al. Interpreting change scores for pain and functional status in low back pain. Spine 2008;33(1):90-94. [doi: 10.1097/brs.0b013e31815e3a10]
53. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
54. Traeger AC, Hübscher M, Henschke N, Moseley GL, Lee H, McAuley JH. Effect of primary care-based education on reassurance in patients with acute low back pain: systematic review and meta-analysis. JAMA Intern Med 2015 May;175(5):733-743. [doi: 10.1001/jamainternmed.2015.0217] [Medline: 25799308]
55. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
56. Cherkin D, Balderson B, Wellman R, Hsu C, Sherman KJ, Evers SC, et al. Effect of low back pain risk-stratification strategy on patient outcomes and care processes: the MATCH randomized trial in primary care. J Gen Intern Med 2018 Aug;33(8):1324-1336 [FREE Full text] [doi: 10.1007/s11606-018-4468-9] [Medline: 29790073]
57. Delitto A, Patterson CG, Stevans JM, Freburger JK, Khoja SS, Schneider MJ, et al. Stratified care to prevent chronic low back pain in high-risk patients: the TARGET trial. A multi-site pragmatic cluster randomized trial. EClinicalMedicine 2021 Apr;34:100795 [FREE Full text] [doi: 10.1016/j.eclinm.2021.100795] [Medline: 33870150]
58. Kloek CJ, Janssen J, Veenhof C. Development of a checklist to assist physiotherapists in determination of patients' suitability for a blended treatment. Telemed J E Health 2020 Aug;26(8):1051-1065. [doi: 10.1089/tmj.2019.0143] [Medline: 31804904]
59. Jones CM, Shaheed CA, Ferreira GE, Kharel P, Christine Lin C, Maher CG. Advice and education provide small short-term improvements in pain and disability in people with non-specific spinal pain: a systematic review. J Physiother 2021 Oct;67(4):263-270 [FREE Full text] [doi: 10.1016/j.jphys.2021.08.014] [Medline: 34518145]
60. Bier JD, Sandee-Geurts JJ, Ostelo RW, Koes BW, Verhagen AP. Can primary care for back and/or neck pain in the Netherlands benefit from stratification for risk groups according to the STarT back tool classification? Arch Phys Med Rehabil 2018 Jan;99(1):65-71. [doi: 10.1016/j.apmr.2017.06.011] [Medline: 28709881]
61. Deyo RA, Mirza SK, Turner JA, Martin BI. Overtreating chronic back pain: time to back off? J Am Board Fam Med 2009;22(1):62-68 [FREE Full text] [doi: 10.3122/jabfm.2009.01.080102] [Medline: 19124635]

62. Williams CM, Maher CG, Latimer J, McLachlan AJ, Hancock MJ, Day RO, et al. Efficacy of paracetamol for acute low-back pain: a double-blind, randomised controlled trial. The Lancet 2014 Nov;384(9954):1586-1596. [doi: 10.1016/s0140-6736(14)60805-9]
63. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011 Dec 05;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
64. Fergusson D, Aaron S, Guyatt G, Hébert P. Post-randomisation exclusions: the intention to treat principle and excluding patients from analysis. BMJ 2002 Sep 21;325(7365):652-654 [FREE Full text] [doi: 10.1136/bmj.325.7365.652] [Medline: 12242181]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
LBP: low back pain
LMM: linear mixed model
MD: mean difference
ODI: Oswestry Disability Index

Edited by A Mavragani; submitted 30.06.21; peer-reviewed by M Merolli, A Martin-Pintado Zugasti, M Salimi, P Pednekar; comments to author 12.08.21; revised version received 23.12.21; accepted 28.12.21; published 25.02.22.

Please cite as:
Koppenaal T, Pisters MF, Kloek CJJ, Arensman RM, Ostelo RWJG, Veenhof C
The 3-Month Effectiveness of a Stratified Blended Physiotherapy Intervention in Patients With Nonspecific Low Back Pain: Cluster Randomized Controlled Trial
J Med Internet Res 2022;24(2):e31675
URL: https://www.jmir.org/2022/2/e31675
doi:10.2196/31675
PMID:35212635

Original Paper

Long-term Effectiveness of a Smartphone App Combined With a Smart Band on Weight Loss, Physical Activity, and Caloric Intake in a Population With Overweight and Obesity (Evident 3 Study): Randomized Controlled Trial

Cristina Lugones-Sanchez[1], NP; Jose I Recio-Rodriguez[1,2], PhD; Cristina Agudo-Conde[1], MSc; Irene Repiso-Gento[3], MD; Esther G Adalia[4], MSc; José Ignacio Ramirez-Manent[5,6], MD, PhD; Maria Antonia Sanchez-Calavera[7,8], MD, PhD; Emiliano Rodriguez-Sanchez[1,9], MD, PhD; Manuel A Gomez-Marcos[1,9*], MD, PhD; Luis Garcia-Ortiz[1,10*], MD, PhD; EVIDENT 3 Investigators[11*]

[1]Primary Care Research Unit of Salamanca (APISAL), Institute of Biomedical Research of Salamanca, Health Service of Castilla y León, Salamanca, Spain

[2]Department of Nursing and Physiotherapy, University of Salamanca, Salamanca, Spain

[3]Renedo de Esgueva Health Center, Health Service of Castilla y León, Valladolid, Spain

[4]Health and Social Research Center, University of Castilla-La Mancha, Cuenca, Spain

[5]Calvià Primary Care Center, The Health Research Institute of the Balearic Islands, Health Service of Balearic Islands, Palma de Mallorca, Spain

[6]Department of Medicine, University of the Balearic Islands, Palma de Mallorca, Spain

[7]Las Fuentes Norte Health Center, Aragonese Group of Primary Care Research (GAIAP), Aragon Health Research Institute (IISA), Aragon Health Service, Zaragoza, Spain

[8]Department of Internal Medicine, Psychiatry and Dermatology, University of Zaragoza, Zaragoza, Spain

[9]Department of Medicine, University of Salamanca, Salamanca, Spain

[10]Department of Biomedical and Diagnostic Sciences, University of Salamanca, Salamanca, Spain

[11]See Acknowledgements, Barcelona, Spain

[*]these authors contributed equally

Corresponding Author:
Cristina Lugones-Sanchez, NP
Primary Care Research Unit of Salamanca (APISAL)
Institute of Biomedical Research of Salamanca
Health Service of Castilla y León
Avda Portugal 83, 2nd Fl.
Salamanca, 37005
Spain
Phone: 34 923291100 ext 54750
Email: crislugsa@gmail.com

Abstract

Background: Multicomponent mobile health approaches can improve lifestyle intervention results, although little is known about their long-term effectiveness.

Objective: This study aims to evaluate the long-term effectiveness (12 months) of a multicomponent mobile health intervention—combining a smartphone app, an activity tracker wristband, and brief counseling, compared with a brief counseling group only—on weight loss and improving body composition, physical activity, and caloric intake in Spanish sedentary adults with overweight or obesity.

Methods: We conducted a randomized controlled, multicenter clinical trial (Evident 3). A total of 650 participants were recruited from 5 primary care centers, with 318 participants in the intervention group (IG) and 332 in the control group (CG). All participants were briefly counseled about a healthy diet and physical activity at the baseline visit. For the 3-month intervention period, the IG received training to use the app to promote healthy lifestyles and the smart band (Mi Band 2, Xiaomi). All measurements were performed at baseline and at 3 and 12 months. Physical activity was measured using the International Physical Activity

Questionnaire–Short Form. Nutritional habits were assessed using the Food Frequency Questionnaire and Adherence to Mediterranean diet questionnaire.

Results: Of the 650 participants included, 563 (86.6%) completed the 3-month visit and 443 (68.2%) completed the 12-month visit. After 12 months, the IG showed net differences in weight (−0.26, 95% CI −1.21 to 0.70 kg; *P*=.02), BMI (−0.06, 95% CI −0.41 to 0.28 points; *P*=.01), waist-height ratio (−0.25, 95% CI −0.94 to 0.44; *P*=.03), body adiposity index (−0.33, 95% CI −0.77 to 0.11; *P*=.03), waist circumference (−0.48, 95% CI −1.62 to 0.66 cm, *P*=.04) and hip circumference (−0.69, 95% CI –1.62 to 0.25 cm; *P*=.03). Both groups lowered daily caloric intake and increased adherence to the Mediterranean diet, with no differences between the groups. The IG increased light physical activity time (32.6, 95% CI −30.3 to 95.04 min/week; *P*=.02) compared with the CG. Analyses by subgroup showed changes in body composition variables in women, people aged >50 years, and married people.

Conclusions: The low-intensity intervention of the Evident 3 study showed, in the IG, benefits in weight loss, some body composition variables, and time spent in light physical activity compared with the CG at 3 months, but once the devices were collected, the downward trend was not maintained at the 12-month follow-up. No differences in nutritional outcomes were observed between the groups.

Trial Registration: ClinicalTrials.gov NCT03175614; https://clinicaltrials.gov/ct2/show/NCT03175614

International Registered Report Identifier (IRRID): RR2-10.1097/MD.0000000000009633

(J Med Internet Res 2022;24(2):e30416) doi:10.2196/30416

KEYWORDS

mobile app; telemedicine; eHealth; weight control; exercise; obesity; mobile phone

Introduction

Background

The combination of excess body weight and physical inactivity contributes to global mortality [1,2] and to a shorter healthy life expectancy [3]. Moreover, both are associated with a major risk for serious chronic diseases, including type 2 diabetes and cardiovascular diseases, as well as increasing cardiovascular risk factors [4]. Such conditions may impact individual quality of life and well-being while increasing the burden on the health system [5]. Several strategies have been implemented to tackle obesity, mainly focusing on changing lifestyles. In general, weight loss interventions aim to increase physical activity, reduce daily energy intake, improve diet or nutritional habits, and achieve psychological changes [6]. Owing to the complex nature of obesity, multicomponent interventions, capable of addressing various aspects related to its causes, are shown to be more effective in reducing cardiovascular risk factors [7] and body weight [8]. However, finding accessible and cost-effective multicomponent strategies that promote healthy lifestyles over time is challenging.

Mobile Health

Mobile health (mHealth) approaches, which are defined as the use of mobile wireless technologies for health [9], could optimize these efforts as portable and flexible tools, as well as improve the follow-up and feedback by registering health information [10] and provide efficient health management assistance for patients [11]. Specifically, some reviews suggested that mHealth could be more effective in losing weight than traditional approaches [12,13]. Among mHealth tools, smartphones are positioned as the most effective approach to achieve weight management [14], showing the most beneficial effects in the short term [15] (≤6 months). However, mobile phone intervention reports modest improvements in other lifestyles such as physical activity [16] or changes in biomarkers [13], highlighting the need to complement the intervention with other tools. Along these lines, wearable devices have garnered attention in improving physical activity and reducing sedentary lifestyle. Pedometer-based interventions have been widely explored [17], but the constant evolution of these tools requires the inclusion of emerging electronic devices [18]. Activity tracker wristbands, also called “smart bands,” have shown their validity and reliability in measuring physical activity outcomes [19] (eg, steps, kilometer walked, and intensity). Furthermore, a recent systematic review found that these device-based interventions are effective for increasing physical activity among chronic disease populations [20], making activity tracker wristbands a good option for inclusion in lifestyle interventions, as wearable activity trackers have the potential to increase physical activity participation [21]. However, wearable activity trackers alone may not be sufficient to achieve the expected lifestyle changes [22], so their inclusion in mHealth multicomponent obesity interventions, which appears to be more effective than app interventions alone [23], could be a beneficial strategy to obtain positive results in diet, physical activity, or other health variables [24,25]. In addition, self-monitoring in digital health interventions is associated with greater weight loss [26], so the use of both approaches could produce better weight outcomes.

Despite these promising results, the evidence for long-term efficacy is still limited, revealing that more evidence of effectiveness over long follow-up periods is required [27]. A systematic review showed that wearable devices might improve long-term physical activity and weight loss outcomes, although a comparison with traditional methods did not show clear benefits [28]. A similar situation exists with regard to multicomponent interventions with wearables, which may provide a strategy to improve long-term weight loss [29] despite the limited data on its effectiveness.

Objectives

Previously, the Evident 2 study evaluated the effect of adding an app (Evident 2) to a standardized intervention designed to improve adherence to the Mediterranean diet and increase physical activity in the general population [30]. The Evident 3 study aimed to evaluate the long-term effectiveness (12 months) of a multicomponent mHealth intervention—combining a later version of the smartphone app (Evident 3) and a smart band, compared with a brief counseling–only group—on body weight loss, improving body composition, increasing time spent in physical activity, and decreasing caloric intake in Spanish sedentary adults with overweight and obesity. The short-term study results (3 months) related to body composition have already been published [31].

Methods

Design and Setting

A multicenter, randomized controlled clinical trial with 2 parallel groups was conducted in a primary care setting (Evident 3 study). The Primary Care Research Unit of Salamanca (APISAL) at the Biomedical Research Institute of Salamanca (IBSAL) coordinated the project in 5 health centers located in Spain from the Network for Preventive Activity and Health Promotion (REDIAPP; Salamanca, Valladolid, Cuenca, Palma de Mallorca, and Zaragoza). The main aim of the study was to evaluate the effect of the intervention on weight loss in participants with overweight and obesity [32] (ClinicalTrials.gov NCT03175614). Between June 2017 and June 2020, evaluations were made at baseline and after the completion of 3 and 12 months. The results presented in this paper correspond to the long-term effect (12 months) of the 3-month mHealth intervention on the primary outcome (weight loss) and secondary outcomes related to physical activity and diet.

Study Population

The participants were selected by random sampling among patients who attended a consultation with their primary care provider in each participating center. The inclusion criteria were age between 20 and 65 years, a BMI between 27.5 kg/m^2 and 40 kg/m^2, classified as sedentary (20 minutes of vigorous-intensity activity ≤3 times per week; 30 minutes of moderate-intensity activity ≤5 times per week; or any combination of moderate and vigorous activity ≤5 times per week [33]), agreement to participate in the study, and signing the informed consent document. A detailed description of the exclusion criteria can be found elsewhere [32].

Sample Size

The sample size calculation was performed for the primary study endpoint of weight loss at 12 months. Accepting an α risk of .05 and a β risk of .20, with an SD of 12 kg, estimated in participants from the Evident 2 study [30], 592 participants would be needed (296 per group) to detect a decrease in weight of ≥3 kg [34] in the intervention group (IG) versus the control group (CG), considering a 15% loss to follow-up. This effect size represents a 3% to 5% difference between groups, which should produce clinically relevant health benefits [35].

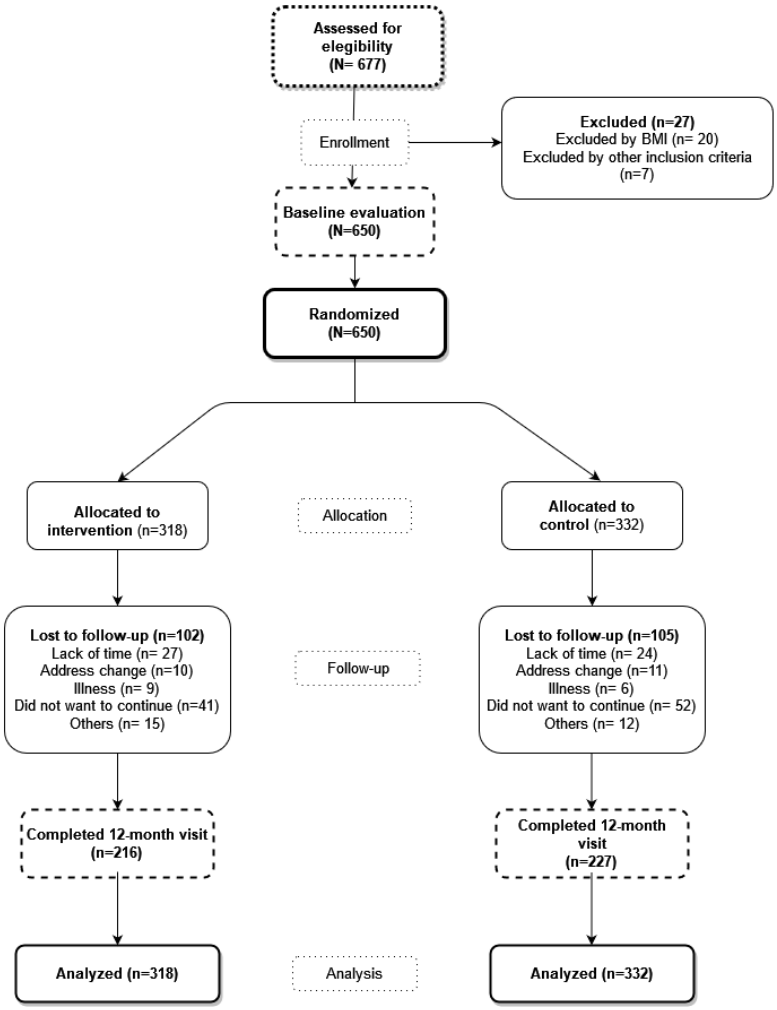
Randomization

Participants were randomly assigned to either the IG or the CG after the baseline visit and provided informed consent. The allocation sequence was generated in a 1:1 ratio using the Epidat (version 4.2; Xunta de Galicia) software package [36] by an independent researcher and concealed until the group was assigned. Owing to the nature of the study, the intervention could not be blinded to the participants.

Procedures

Each participant completed an initial visit and 2 follow-up visits at 3 and 12 months after randomization (Figure 1). They did not receive any compensation for the study or visit completion. Data from the visits were collected by a research nurse on a paper-based Case Report Form and recorded after the visit on the study website. The IG completed an additional appointment 7 days from the baseline visit to explain the use of the smartphone app and set it with the participant's data. The researcher who carried out the additional appointment was different from the researcher who collected the data of the visits.

Figure 1. Flowchart of the Evident 3 study.



Primary and Secondary Outcomes

Overview

The primary outcome was weight loss (kg). The secondary outcomes included changes in physical activity (min/week), caloric intake (kcal/day), adherence to the Mediterranean diet (points), and changes in body composition (anthropometric indexes). All outcomes were measured at baseline and at 3 and 12 months after randomization.

Weight Loss

Body weight was measured twice to the nearest 0.1 kg, with the participant barefoot and wearing light clothing, using a homologated electronic balance (Scale 7830; Soehnle Professional GmbH & Co). Height was measured with the participant barefoot in the standing position using a portable system (Seca 222; Medical scale and measurement system). BMI was calculated by dividing weight (kg) by height squared (m^2). Waist circumference was measured twice on bare skin, using a flexible tape parallel to the floor, at the level of the upper border of the iliac crest, with the participant standing and after inspiration, following the recommendations of the Spanish Society for the Study of Obesity [37]. Hip circumference was similarly measured at the level of the trochanters over the underwear.

In addition, anthropometric indices were estimated to evaluate body composition changes:

Waist–height ratio =

$$WC / Height \quad \textbf{(1)}$$

Body adiposity index =

$$[HC / (Height^{1.5})] - 18 \quad \textbf{(2)}$$

Waist–hip ratio =

$$WC / HC \quad \textbf{(3)}$$

Body shape index [38] =

$$WC / [(BMI^{0.666}) \times (Height^{0.5})] \quad \textbf{(4)}$$

Body roundness index [39] =

$$364.2 - 365.5 \times \sqrt{\{1 - [WC / (2 \times 3.1416)]^2 / (0.5 \times Height)^2\}} \quad \textbf{(5)}$$

Physical Activity

Physical activity was assessed using the short version of the International Physical Activity Questionnaire–Short Form (IPAQ-SF) [40]. The IPAQ-SF is a self-reported questionnaire that evaluates sitting and active time in the last 7 days through 7 questions. The categories consisted of walking, moderate-intensity activity, and vigorous-intensity activity, according to the energy expenditure estimated for each of them:

3.3, 4.0, and 8.0 metabolic equivalents (METs), respectively; 1 MET being the amount of oxygen consumed while sitting at rest [41]. Thus, the IPAQ-SF enables METs-minutes per week to be calculated, and participants were stratified into 3 activity levels (low, intermediate, and high).

Nutritional Habits

Caloric intake (kcal/day) and dietary habits of participants were evaluated using a semiquantitative Food Frequency Questionnaire previously validated in the Spanish population [42]. The frequency options are divided into 9 intake categories, ranging from never to >6 servings per day. Food Frequency Questionnaire data were used to estimate the daily intake of macro- and micronutrients and the mean kcal/day.

Adherence to the Mediterranean diet was assessed using the validated 14-point Mediterranean Diet Adherence Screener [43] developed by the prevention with Mediterranean diet study group, which comprises 12 questions on food consumption frequency and 2 questions on food intake habits. Each question was scored as 0 or 1, and the total score ranged from 0 to 14. A total score of 9 points or higher indicated adequate adherence.

Adherence to Self-monitoring on the Smartphone App

Adherence to self-monitoring on the smartphone app was assessed by the number of days that the participants logged into the app and recorded any dish or food. Records were classified into 4 categories: 0 days, 1 to 30 days, 31 to 60 days, and >60 days. Participants who used the app for >60 days during the 3 months that they had the app were classified as sufficiently adherent, whereas ≤60 days of use was classified as having low adherence.

Other Variables

Sociodemographic Variables

Data on age, sex, marital status, educational level, and occupation were collected at the time of inclusion in the study.

Peripheral Blood Pressure

Three measurements of systolic and diastolic blood pressure were performed using the average of the last 2 measurements with a validated Omron M10-IT model sphygmomanometer (Omron Healthcare). The measurements were made on both arms, with the participant seated, after at least 5 minutes of rest with an appropriately sized cuff, following the recommendations of the European Society of Hypertension [44].

Smoking Status

This was assessed through a questionnaire of 4 standard questions adapted from the World Health Organization monitoring of trends and determinants in cardiovascular disease study [45]. Study participants were classified as current smokers, former smokers (>1 year without smoking), or nonsmokers.

Intervention

A detailed description of brief counseling and specific intervention has been published in the study protocol [32]. All the intervention materials were provided in Spanish.

Standard Counseling (CG and IG)

A trained nurse at each primary health center, who was not involved in other aspects of the study, gave 5 minutes of lifestyle counseling to both groups (CG and IG) before randomization, focusing on physical activity and diet in compliance with the international recommendations for the general population. The health benefits of physical activity were explained as well as the recommendation to complete at least 30 minutes of moderate activity 5 days a week, or 20 minutes of vigorous activity 3 days a week. Counseling on food was in compliance with the plate method [46], in which a plate is divided into 4 parts: half the plate for salad or vegetables, one-quarter for proteins (white meat preferred over red meat), and the final quarter for carbohydrates. In addition, a medium-sized piece of fruit and a skimmed dairy product should be consumed for dessert. This advice enhanced the intake of healthy food, according to the Mediterranean diet pattern, and daily caloric intake goals were not included. No reinforcement of counseling was offered at any other study visit or between the 3- and 12-month visits.

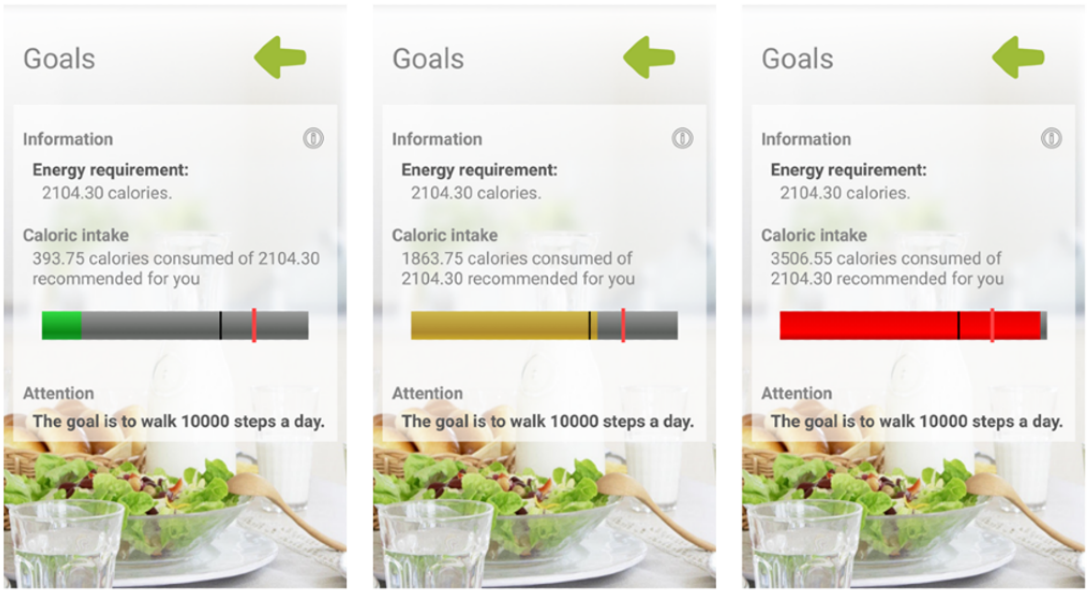
Specific Intervention (IG)

The IG received a low-intensity intervention consisting of a smartphone with the EVIDENT 3 app (Samsung Galaxy J3) and a smart band (Xiaomi Mi Band 2) for 3 months, corresponding to the length of the intervention, without any additional reinforcement or counseling by the investigators throughout the study. Participants were trained at another 15-minute visit scheduled 7 days after the baseline visit in the use of the app (EVIDENT 3 app [record entry no. 00/2017/2438]) specifically designed for the study by CGB Computer Company and APISAL, as well as the use of the smart band, instructing them to use both tools daily.

During this visit, the app was configured with each participant's data (sex, age, weight, and height). It was designed to allow full daily self-monitoring of food intake (Figure 2) and automatically record physical activity through the smart band, which was configured to synchronize with the app. Participants entered their food intake daily by selecting dishes and foods from the app menu and indicating the portion size. Food composition data were collected from the Spanish Food Composition Database [47], developed by Spanish Food Composition Database Network and Spanish Agency for Food Safety and Nutrition. Once all the daily information is collected, the app integrates the data to create personalized healthy food recommendations based on the Mediterranean diet pattern and specific targets for daily calorie intake that would lead to weight loss. The app displays the amount of calories recorded (Figure 2) and a bar that changes color (green, yellow, or red) according to the level set. It was configured to achieve a hypocaloric diet, calculating the upper limit (the red line) by adding, according to age and sex, the basal metabolic rate, diet-induced thermogenesis, and estimated energy expenditure for sedentary activities. The lower limit (the black line) was 85% of the calories calculated, and below this, the bar appeared in green; between the red and black lines, it appeared in yellow; and above the red line, it appeared in red. The participant was able to consult the app for these recommendations as well as information about caloric intake changes and macronutrient

distribution (carbohydrates, proteins, and unsaturated and saturated fats). The smart band was set to congratulate the user when reaching 10,000 steps/day, and the app displayed this step recommendation in the goals section. Behavioral strategies were included in the mHealth intervention to enhance self-efficacy using self-monitoring, goal-setting, and positive reinforcement. At the 3-month visit, participants returned the intervention tools to the researchers. Thereafter, participants did not have access to the intervention devices and were advised not to use other digital tools for weight loss until the end of the study. All the information generated by the app was duly analyzed and entered into the database. In addition, once the tools were returned, monthly mean daily steps and activity minutes were collected for the last 2 of the 3 months of the intervention from the Mi Fit app (Xiaomi) to assess whether the smart band was worn.

Figure 2. Evident 3 app screenshots.



Blinding Strategy

The researcher who carried out the specific intervention was different from the person responsible for the assessment and the standard counseling; both were kept blinded throughout the study, as was the investigator who conducted the data analysis. Owing to the nature of the study, the participants could not be blinded. To prevent contamination between groups, in the follow-up visits (3 and 12 months), only the study variables were evaluated, but no advice or reinforcement was provided. In addition, the app was not available for download on the internet or anywhere until the end of the study, so the CG was not able to make use of it in any way. During the follow-up visits, participants were instructed not to use other digital health technologies.

Statistical Analysis

Baseline characteristics of the study population were expressed as mean and SD for quantitative variables and as frequency distributions for categorical variables. Student *t* test (2-tailed), chi-square test, and Fisher exact test were used to determine differences in baseline characteristics between the IG and CG. Analyses of the results were performed on an intention-to-treat basis. Paired Student *t* test or McNemar test was applied to assess changes within the same group. To analyze the effect of the intervention, in the follow-up for primary and secondary endpoints, we performed several multivariate analyses of the variance of repeated measures using the general linear model, comparing the changes observed between the IG and CG in the analyzed variables, first unadjusted and then adjusted for age and sex.

We performed a subanalysis by multivariate analysis of the variance of repeated measures of the intervention effect in primary and secondary outcomes stratified by sex (men and women), age (<50 years and ≥50 years), and marital status (married, single, or others).

The contrast in the hypotheses established an α value of .05. The data were analyzed using SPSS Statistics software (version 26.0; IBM Corporation).

Ethical Considerations

The study was approved by the clinical research ethics committee of the Health Area of Salamanca in April 2016. In addition, the study was approved by the ethics committees of the 4 collaborating centers: Aragón, Castilla-la Mancha, Baleares, and Valladolid Oeste. All procedures were performed in accordance with the ethical standards of the institutional research committee and the 2013 Declaration of Helsinki [48]. All participants signed written informed consent documents before participation in the study. The trial was registered at ClinicalTrials.gov with identifier NCT03175614 on May 31, 2017.

Results

Baseline Characteristics of the Participants and Follow-up

Of the 650 participants who completed the baseline visit, 563 (86.6%) completed the 3-month visit and 443 (68.2%) completed the 12-month visit. There were 207 (207/650, 31.8%) participants who dropped out of the study, 32.1% (102/318) in the IG and 31.6% (105/332) in the CG. Participants assigned to each group and the reasons for withdrawal from the trial may be consulted in Figure 1.

The clinical and sociodemographic baseline characteristics of the 650 participants are presented in Table 1. The mean age of the entire sample was 48.31 (SD 9.67) years, with a mean BMI of 33.0 (SD 3.48) kg/m^2. In addition, 68.5% (445/650) of participants were women, 68.3% (444/650) were married, and 46.8% (304/650) were aged ≥50 years. No differences were observed between the study groups at baseline.

The comparison between the baseline characteristics of the 207 participants who dropped out and those who completed the study are shown in Multimedia Appendix 1. It should be noted that those who dropped out were younger (46.3 vs 49.2 years) and had a higher weight (93.3 vs 90.2 kg) and BMI (33.6 vs 32.7 kg/m^2), with no difference in the rest of the variables analyzed.

Table 1. Baseline characteristics of study participants.

Characteristics	Intervention (n=318)	Control (n=332)
Age (years)		
Value, mean (SD)	47.7 (10.1)	48.9 (9.2)
<50, n (%)	178 (56)	168 (50.6)
>50, n (%)	140 (44)	164 (49.4)
Sex, n (%)		
Men	104 (32.7)	101 (30.4)
Women	214 (67.3)	231 (69.6)
Marital status, n (%)		
Single	60 (18.9)	74 (22.3)
Married	222 (69.8)	222 (66.9)
Separated	31 (9.7)	30 (9)
Widower	5 (1.6)	6 (1.8)
Employment status, n (%)		
Works outside of home	232 (72.9)	249 (75.1)
Homemaker	22 (6.9)	21 (6.3)
Retired	22 (6.9)	19 (5.7)
Student	9 (2.8)	5 (1.5)
Unemployed	33 (10.4)	38 (11.4)
Educational level, n (%)		
University studies	122 (38.5)	134 (40.5)
Middle or high school	158 (48.9)	152 (45.9)
Elementary school	37 (11.7)	45 (13.6)
Clinical variables, mean (SD)		
Weight (kg)	91.4 (14.8)	91.1 (14.8)
BMI (kg/m^2)	33.1 (3.4)	33.0 (3.6)
Waist circumference (cm)	107.4 (12.9)	107.4 (10.7)
Systolic blood pressure (mm Hg)	119 (15)	120 (16)
Diastolic blood pressure (mm Hg)	79 (9)	81 (10)
Heart ratio (bpm[a])	72 (12)	74 (12)
Total cholesterol (mg/dL)	198 (36)	202 (40)
HDL[b] cholesterol (mg/dL)	51 (13)	52 (12)
BMI classification (kg/m^2), n (%)		
27.5-29.9	75 (23.6)	82 (24.7)
30-40	243 (76.4)	250 (75.3)
Chronic diseases, n (%)		
Hypertension	88 (27.7)	116 (35.0)
Dyslipidemia	73 (23.4)	87 (26.5)
Diabetes mellitus	5 (1.7)	4 (1.3)
Medication use, n (%)		
Antihypertensive drugs	50 (15.7)	69 (20.8)
Lipid lowering drugs	50 (15.7)	56 (16.9)

Characteristics	Intervention (n=318)	Control (n=332)
Hypothyroid drugs	31 (9.8)	37 (11.1)

[a]bpm: beats per minute.

[b]HDL: high-density lipoprotein.

Adherence to Self-monitoring on the Smartphone App

Adherence to self-monitoring on the smartphone app was calculated from the app output data by an independent researcher. The median app use was 64.5 days out of the 90 days of the intervention (71.67%). Of the 318 participants assigned to the IG, 150 (47.2%) adhered sufficiently by recording data in the app between 61 and 90 days. In total, 3 participants did not register any food, and there were 36 data files from which no information was available for technical reasons (Multimedia Appendix 2). Multimedia Appendix 3 displays the median days of app use out of the 90 days in percentage to show adherence to self-monitoring on the app, grouped by sex, marital status, and age.

Changes in Weight and Anthropometric Variables During the Study Period

Table 2 shows the decrease in body weight at 3 and 12 months in both groups. Comparing groups, the IG achieved a net weight loss difference of 0.76 (95% CI −1.33 to −0.19) kg at 3 months and 0.26 (95% CI −1.21 to 0.70; *P*=.02) kg at 12 months, more than the CG. The overall weight reduction of the IG was 2.05% at 3 months and 1.58% at 12 months, while the CG showed 1.1% and 1.26% reductions in body weight at 3 and 12 months, respectively. Only 18.2% (52/285) of the IG participants achieved a clinically significant weight loss of ≥5% at the 3 months visit and 19.4% (42/216) of them achieved that percentage at 12 months. In the CG, 12.9% (35/271) achieved a weight loss of 5% at 3 months and 18.5% (42/227) reached that loss at the 12-month visit. Regarding adherence to self-monitoring on the app, both in number of days used and median percentage of days, a positive correlation was found at 3 months with weight loss (*r*=0.239; *P*<.001) and BMI (*r*=0.203; *P*<.001) but not at 12 months or with weight (*r*=0.015; *P*=.83) and BMI (*r*=0.021; *P*=.77).

In addition, the IG showed changes in waist circumference (−0.76, 95% CI −1.47 to −0.05 cm) and hip circumference (−1.02, 95% CI −1.68 to −0.34 cm) compared with the CG at 3 months. Similar results were found at 12 months, with net decreases in waist circumference (−0.48, 95% CI −1.62 to 0.66 cm) and hip circumference (−0.69, 95% CI −1.62 to 0.25 cm; *P*=.04 and *P*=.03, respectively) between groups. Regarding body composition parameters, BMI decreased at 3 months (−0.30, 95% CI −0.52 to −0.09 points) and slightly changed at 12 months (−0.06, 95% CI −0.41 to 0.28 points; *P*=.01), comparing the study groups. Similar results were found for waist–height ratio at 3 months (−0.48, 95% CI −0.92 to −0.04) and 12 months (−0.25, 95% CI −0.94 to 0.44; *P*=.03) and body adiposity index at 3 months (−0.50, 95% CI −0.81 to −0.18) and 12-month follow-up (−0.33, 95% CI −0.77 to 0.11; *P*=.03) between groups. Although the IG tended to show decreases in the rest of the indexes analyzed (waist–hip ratio, body shape index, and body roundness index) at both follow-up visits, no significant differences were observed between groups.

Figure 3 shows the evolution of the main anthropometric parameters analyzed over time, with a decrease at 3 months in both groups, especially in the IG. However, the downward trend was not maintained in the IG at 12 months, while the CG continued to decrease.

Table 2. Effect of the mobile health intervention on body weight and other anthropometric parameters.

Parameters	Intervention group (n=318)		Control group (n=332)		Net difference		
	Values	*P* value	Values	*P* value	Values	*P* value[a]	*P* value[b]
Weight (kg)							
Baseline, mean (SD)	91.4 (14.8)	N/A[c]	91.1 (14.8)	N/A	N/A	N/A	N/A
3-month change, mean difference (95% CI)	−1.79 (−2.20 to −1.37)	<.001	−1.03 (−1.41 to −0.64)	<.001	−0.76 (−1.33 to −0.19)	N/A	N/A
12-month change, mean difference (95% CI)	−1.46 (−2.15 to −0.77)	<.001	−1.20 (−1.87 to −0.54)	<.001	−0.26 (−1.21 to 0.70)	.03	.02
Waist circumference (cm)							
Baseline, mean (SD)	107.4 (12.9)	N/A	107.4 (10.7)	N/A	N/A	N/A	N/A
3-month change, mean difference (95% CI)	−2.18 (−2.71 to −1.65)	<.001	−1.42 (−1.90 to −0.94)	<.001	−0.76 (−1.47 to −0.05)	N/A	N/A
12-month change, mean difference (95% CI)	−2.28 (−3.14 to −1.43)	<.001	−1.80 (−2.57 to −1.04)	<.001	−0.48 (−1.62 to 0.66)	.04	.04
Hip circumference (cm)							
Baseline, mean (SD)	116.4 (11.4)	N/A	115.5 (9.3)	N/A	N/A	N/A	N/A
3-month change, mean difference (95% CI)	−1.96 (−2.42 to −1.50)	<.001	−0.94 (−1.42 to −0.47)	<.001	−1.02 (−1.67 to −0.36)	N/A	N/A
12-month change, mean difference (95% CI)	−1.81 (−2.47 to −1.16)	<.001	−1.13 (−1.79 to −0.46)	.001	−0.69 (−1.62 to 0.25)	.03	.03
BMI (kg/m^2)							
Baseline, mean (SD)	33.1 (3.4)	N/A	32.9 (3.6)	N/A	N/A	N/A	N/A
3-month change, mean difference (95% CI)	−0.69 (−0.85 to −0.53)	<.001	−0.38 (−0.52 to −0.24)	<.001	−0.30 (−0.52 to −0.09)	N/A	N/A
12-month change, mean difference (95% CI)	−0.49 (−0.74 to −0.24)	<.001	−0.43 (−0.66 to −0.19)	<.001	−0.06 (−0.41 to 0.28)	.02	.01
Waist–height ratio							
Baseline, mean (SD)	64.82 (6.9)	N/A	64.76 (5.7)	N/A	N/A	N/A	N/A
3-month change, mean difference (95% CI)	−1.35 (−1.68 to −1.02)	<.001	−0.87 (−1.16 to −0.58)	<.001	−0.48 (−0.92 to −0.04)	N/A	N/A
12-month change, mean difference (95% CI)	−1.34 (−1.86 to −0.83)	<.001	−1.09 (−1.56 to −0.63)	<.001	−0.25 (−0.94 to 0.44)	.04	.03
Body adiposity index							
Baseline, mean (SD)	36.76 (6.6)	N/A	36.31 (5.6)	N/A	N/A	N/A	N/A
3-month change, mean difference (95% CI)	−0.94 (−1.16 to −0.72)	<.001	−0.44 (−0.66 to −0.22)	<.001	−0.50 (−0.81 to −0.18)	N/A	N/A
12-month change, mean difference (95% CI)	−0.85 (−1.16 to −0.54)	<.001	−0.52 (−0.83 to −0.20)	.001	−0.33 (−0.77 to 0.11)	.03	.03
Waist–hip ratio							
Baseline, mean (SD)	0.9 (0.1)	N/A	0.9 (0.1)	N/A	N/A	N/A	N/A
3-month change, mean difference (95% CI)	0.00 (−0.01 to 0.00)	.07	0.00 (−0.01 to 0.00)	.02	0.00 (−0.01 to 0.01)	N/A	N/A
12-month change, mean difference (95% CI)	−0.01 (−0.01 to 0.00)	.13	−0.01 (−0.01 to 0.00)	.03	0.00 (−0.01 to 0.01)	.51	.46
Body shape index[d]							
Baseline, mean (SD)	8.1 (0.6)	N/A	8.1 (0.5)	N/A	N/A	N/A	N/A
3-month change, mean difference (95% CI)	−0.06 (−0.09 to −0.02)	.002	−0.04 (−0.08 to −0.01)	.01	−0.01 (−0.06 to 0.036)	N/A	N/A

Parameters	Intervention group (n=318)		Control group (n=332)		Net difference		
	Values	*P* value	Values	*P* value	Values	*P* value[a]	*P* value[b]
12-month change, mean difference (95% CI)	−0.09 (−0.14 to −0.04)	<.001	−0.06 (−0.11 to −0.01)	.01	−0.03 (−0.09 to 0.043)	.52	.56
Body roundness index							
Baseline, mean (SD)	6.7 (1.9)	N/A	6.6 (1.4)	N/A	N/A	N/A	N/A
3-month change, mean difference (95% CI)	−0.32 (−0.40 to −0.24)	<.001	−0.21 (−0.28 to −0.14)	<.001	−0.11 (−0.21 to 0.00)	N/A	N/A
12-month change, mean difference (95% CI)	−0.31 (−0.43 to −0.18)	<.001	−0.26 (−0.37 to −0.14)	<.001	−0.05 (−0.22 to 0.12)	.07	.05

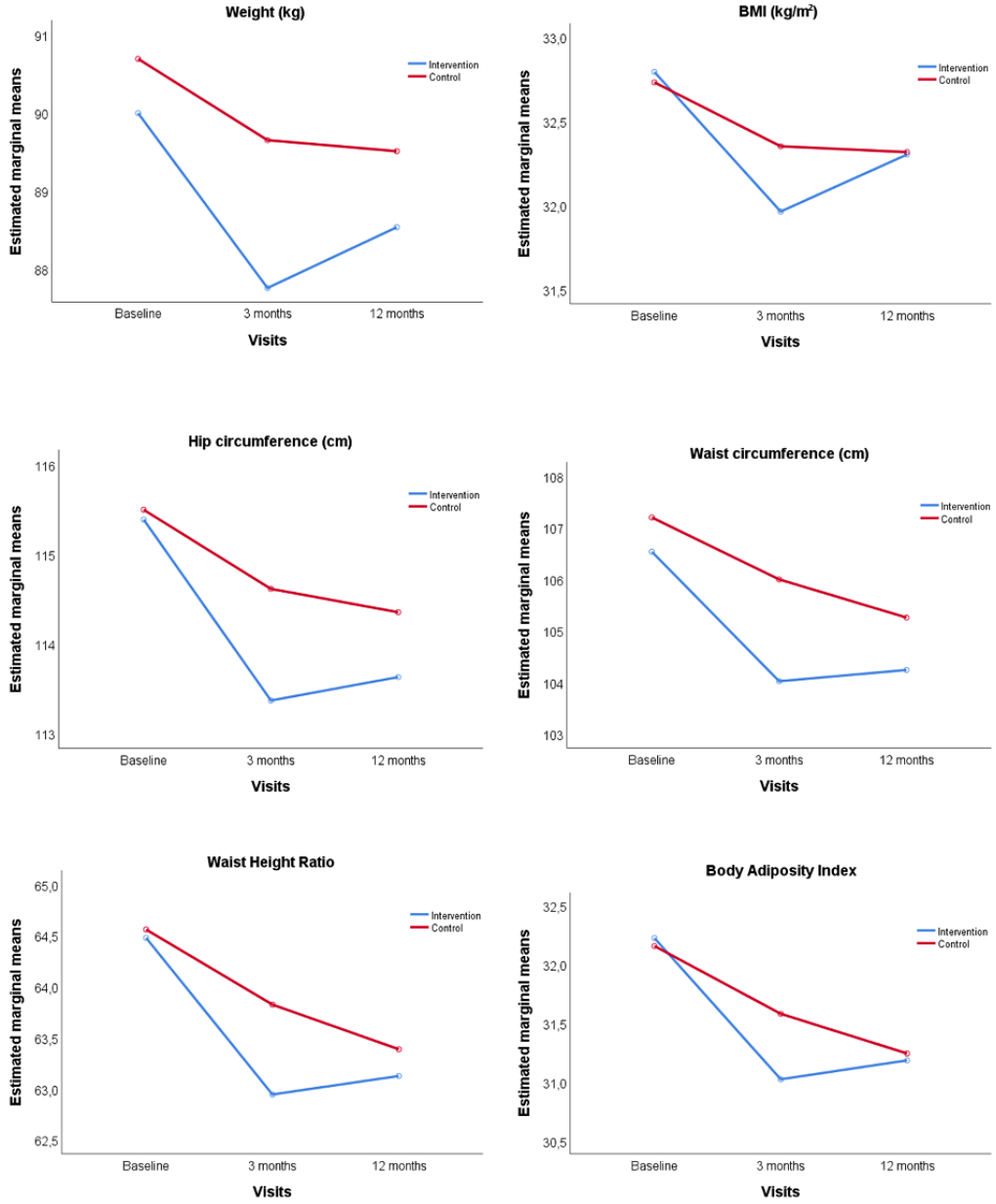
[a]*P* value by analysis of variance.

[b]*P* value by analysis of variance adjusted by age and sex.

[c]N/A: not applicable.

[d]Body Round Shape results are displayed multiplied by 100 for easier reading.

Figure 3. Evolution of weight, BMI, and other anthropometrics parameters from baseline to 3 and 12 months comparing the intervention and control group. *P* value between groups was adjusted by age and sex: weight, *P*=.02; waist circumference, *P*=.04; hip circumference, *P*=.03; BMI, *P*=.01; waist–height ratio, *P*=.03; and body adiposity index, *P*=.03.



Changes in Diet and Physical Activity

In terms of diet, daily caloric intake was lower in both groups (IG: −295, 95% CI −391.63 to −198.93 kcal/day; CG: −222, 95% CI −310.62 to −135.14 kcal/day) at 12 months, with no significant differences between groups (Table 3). The Mediterranean diet adherence increased in both groups at 3 and 12 months, but the differences were not statistically significant. The IG experienced a trend toward increased adherence to the Mediterranean diet at 3 months, but decreased at 12 months, while adherence was maintained in the CG.

Physical activity time in all intensities (light, moderate, and vigorous) assessed by the IPAQ-SF increased in both groups at 3 and 12 months, whereas the time of sedentarism decreased. Although the IG tended to show greater increases in light activity, vigorous activity, and total activity time, only light physical activity (LPA) time showed a net increase of 81.2 (95% CI 31.93-130.58) minutes per week at 3 months and 32.6 (95% CI −30.31 to 95.04) minutes per week at 12 months (*P*=.02) compared with the CG. Regarding the smart band, Multimedia Appendix 4 shows the correlation between the IPAQ-SF data (METs per week and min/week for each activity level) at 3 months in the IG and the information collected from the smart band after 3 months of the intervention (daily average of steps and activity minutes), with a low correlation between variables, which was higher in women than in men.

Table 3. Effect of the mobile health intervention on diet and physical activity variables.

Variables	Intervention group (n=318)		Control group (n=332)		Net difference	
	Value	*P* value	Value	*P* value	Value	*P* value[a]
Energy intake (kcal)						
Baseline, mean (SD)	2394.1 (676.4)	N/A[b]	2359.9 (681.1)	N/A	N/A	N/A
3-month change, mean difference (95% CI)	−204.18 (−285.46 to −122.89)	<.001	−189.84 (−268.94 to −110.73)	<.001	−14.34 (−127.48 to 98.80)	N/A
12-month change, mean difference (95% CI)	−295.28 (−391.63 to −198.93)	<.001	−222.88 (−310.62 to −135.14)	<.001	−72.40 (−202.13 to 57.33)	.59
Score for adherence to MD[c] (points)						
Baseline, mean (SD)	7.1 (1.9)	N/A	7.1 (1.9)	N/A	N/A	N/A
3-month change, mean difference (95% CI)	0.56 (0.34 to 0.78)	<.001	0.41 (0.19 to 0.60)	<.001	0.15 (−0.15 to 0.45)	N/A
12-month change, mean difference (95% CI)	0.37 (0.12 to 0.63)	.005	0.56 (0.29 to 0.82)	<.001	−0.19 (−0.56 to 0.17)	.06
Score adherence to MD (≥9 points)						
Baseline, n (%)	81 (25.3)	N/A	84 (25.4)	N/A	N/A	N/A
3-month change, percentage difference (95% CI)	8.04 (2.07 to 14.02)	.009	5.90 (0.22 to 11.59)	.04	2.14 (−6.11 to 10.39)	N/A
12-month change, percentage difference (95% CI)	6.45 (−0.89 to 13.80)	.09	11.66 (4.74 to 18.58)	.001	−5.21 (−15.26 to 4.85)	.19
Light activity (min/week)						
Baseline, mean (SD)	259.9 (283.4)	N/A	259.8 (287.9)	N/A	N/A	N/A
3-month change, mean difference (95% CI)	102.85 (70.46 to 135.26)	<.001	21.60 (−15.96 to 59.16)	.26	81.26 (31.93 to 130.58)	N/A
12-month change, mean difference (95% CI)	91.21 (58.76 to 125.65)	<.001	58.85 (6.24 to 111.46)	.03	32.36 (−30.31 to 95.04)	.02
Moderate activity (min/week)						
Baseline, mean (SD)	53.0 (154.2)	N/A	41.3 (133.6)	N/A	N/A	N/A
3-month change, mean difference (95% CI)	19.10 (−7.80 to 46.01)	.16	44.32 (22.06 to 66.57)	<.001	−25.21 (−60.27 to 9.85)	N/A
12-month change, mean difference (95% CI)	2.04 (−28.58 to 32.65)	.90	38.90 (13.54 to 64.26)	.003	−36.86 (−76.39 to 2.66)	.06
Vigorous activity (min/week)						
Baseline, mean (SD)	39.2 (228.0)	N/A	30.7 (143.2)	N/A	N/A	N/A
3-month change, mean difference (95% CI)	20.80 (−1.26 to 42.85)	.06	13.73 (−15.81 to 42.26)	.36	7.07 (−29.44 to 43.58)	N/A
12-month change, mean difference (95% CI)	25.29 (2.98 to 47.59)	.03	17.37 (−4.92 to 39.66)	.13	7.92 (−23.54 to 39.37)	.82
Moderate to vigorous activity (min/week)						
Baseline, mean (SD)	92.2 (293.5)	N/A	71.9 (195.4)	N/A	N/A	N/A
3-month change, mean difference (95% CI)	39.89 (5.07 to 74.73)	.03	58.04 (16.83 to 99.26)	.006	−18.15 (−71.77 to 35.48)	N/A
12-month change, mean difference (95% CI)	27.33 (−15.38 to 70.03)	.21	56.27 (18.59 to 93.96)	.004	−28.95 (−85.63 to 27.73)	.33
Total activity time (min/week)						
Baseline, mean (SD)	351.1 (415.8)	N/A	331.7 (345.2)	N/A	N/A	N/A
3-month change, mean difference (95% CI)	142.75 (94.90 to 190.61)	<.001	79.65 (24.27 to 135.02)	.005	63.11 (−9.68 to 135.90)	N/A

Variables	Intervention group (n=318)		Control group (n=332)		Net difference	
	Value	*P* value	Value	*P* value	Value	*P* value[a]
12-month change, mean difference (95% CI)	118.53 (60.59 to 176.48)	<.001	115.12 (45.01 to 185.23)	.001	3.41 (−87.63 to 94.46)	.57
Total sitting time (min/week)						
Baseline, mean (SD)	2903.6 (1397.3)	N/A	2805.9 (1347.6)	N/A	N/A	N/A
3-month change, mean difference (95% CI)	−174.98 (−294.58 to −55.40)	.004	−42.74 (−160.90 to 75.42)	.48	−132.24 (−300.17 to 35.70)	N/A
12-month change, mean difference (95% CI)	−259.84 (−389.54 to −130.16)	<.001	−94.28 (−239.15 to 50.60)	.20	−165.57 (−359.74 to 28.60)	.09

[a]*P* value by analysis of variance.

[b]N/A: not applicable.

[c]MD: Mediterranean diet.

Analyses of the Effect Stratified by Baseline Characteristics

Only women experienced decreases in weight, waist, and other anthropometric parameters analyzed, except for waist–hip ratio, body shape index, and body roundness index, with a net weight loss of 0.50 (95% CI −1.53 to 0.54; *P*=.002) kg at 12 months. With regard to age, those aged <50 years showed a decrease in hip circumference, while people aged >50 years showed a reduction in weight and BMI. Finally, analysis by marital status showed that only married people showed decreases in weight (−0.90, 95% CI −2.0 to 0.2 kg; *P*=.02), BMI (−0.3, 95% CI −0.7 to 0.1 points; *P*=.01), and other anthropometric parameters at 12 months. These analyses are presented in Multimedia Appendix 5.

Discussion

Principal Findings

Overview

The Evident 3 study evaluated the intervention effect and its maintenance over time at 3 and 12 months after the baseline visit. The main findings of the use of a smartphone app in combination with an activity tracker wristband for 3 months and brief counseling showed a greater weight loss compared with the CG, but once these devices were collected, the trend was not maintained at 12 months. Although both groups had reduced weight, BMI, and other anthropometric variables, the IG showed a greater trend to reduce weight, BMI, waist and hip circumference, waist–height ratio, and body adiposity index at the 3-month follow-up and did not maintain a downward trend at the 12-month visit. Regarding diet, both groups decreased their caloric intake (kcal) and improved their Mediterranean diet adherence, but no differences between groups were found. A similar result was observed in physical activity, where physical activity time increased in both groups, but the IG showed an increase in weekly LPA time at 12 months. Analyses stratified by baseline characteristics showed patterns toward greater changes in body composition variables in women, people aged >50 years, and married people.

Weight

Findings from meta-analyses have shown that mHealth weight loss interventions were effective in comparison with minimal intervention or control in the short term but with inconclusive long-term results [15,16]. The Evident 3 study provides insights into the long-term effects of weight loss, BMI, and other anthropometric variables at 3 months, but this trend was not maintained in the period when they no longer had access to the devices up to the 12-month visit. In addition, overall weight reduction in the IG was 2.05% at 3 months, so clinically relevant weight loss was not achieved (>5%).

Despite the heterogeneity found in these types of interventions, some systematic reviews [23,49] suggest that the combination of mHealth tools could be useful for changing lifestyles to healthier ones, but its effect on weight loss remains unclear. Along these lines, the Innovative Approaches to Diet, Exercise and Activity study, which provided a wearable technology combined with a website to the enhanced IG, found no significant difference in weight or physical activity between groups [29] across the 24-month intervention in young adults. Moreover, another study that evaluated the effect of a web-based weight loss program with and without an activity tracking device found that its addition to the intervention did not produce higher changes in weight loss at the 12-month follow-up [50] than those in the CG. In contrast, the Quant study, whose intervention included 3 feedback devices (Bioelectrical Impedance Analysis scale, blood pressure, and step counter) found a positive effect on fat loss at the 12-month follow-up [51]. These results are consistent with those found in our study, where there were differences between groups in weight, body composition variables, and physical activity but without reaching clinically relevant results.

Notwithstanding the effectiveness of self-monitoring behaviors in weight management [52], the challenge lies in finding a way to keep the users using the app, because the frequency of logging into the app is highly related to weight loss success in web-based interventions [53] and mHealth apps [52], as the user adopts new behaviors over time that is supported by the tools. In this regard, the rate of users who adhered sufficiently to the study app (150/318, 47.2%) may be insufficient to show better results in the main outcomes at 12 months. It should also be noted that

between the period in which mHealth tools were collected at the 3-month follow-up and the 12-month visit, there was no reinforcement or any contact with the participants by the researchers. Therefore, despite not obtaining clinically relevant weight loss, the results suggest that the intervention should be modified (longer intervention period, improved adherence strategies, some professional support, etc), as the trend of weight reduction was not maintained over time when tools were removed. Moreover, the CG also reported weight loss but lower than that of the IG, which is commonly reported in weight loss interventions [25], especially if the CG received usual care [54].

Physical Activity and Diet

The study intervention, which included an activity tracker wristband, was shown to increase weekly LPA time measured by the IPAQ-SF. Increasing LPA may improve important health outcomes, such as markers of lipid and glucose metabolism and mortality in the general population [55]. Although the general physical activity recommendations are based on promoting moderate to vigorous physical activity, increasing LPA in the sedentary population may be a good starting point for decreasing inactivity in people with overweight and obesity [56]. Previous studies have reported beneficial changes in physical activity variables, observing a small increase in moderate to vigorous physical activity in women [57] and people with overweight and obesity [22] or in steps per day [20] in people with chronic diseases. Specifically, a recent study reported increased resistance training and reduced energy intake at 6 months using a multicomponent mHealth intervention [25], showing the potential of these tools in physical activity promotion by allowing a more tailored intervention and greater feedback.

Furthermore, energy intake was reduced, and adherence to the Mediterranean diet increased in both groups. However, the app intervention did not achieve better results than the CG. Potential explanations for this include that the brief diet counseling using the plate method was explained at the end of the baseline visit, which could lead to an improvement in the entire sample. Moreover, Solbrig et al [58] suggested that there is a mismatch between people's need for help and what weight management apps provide, as people dislike counting calories, the basis of most self-management apps, and need more tailored support and motivational elements. Along these lines, a recent study found that a digital app that provides personalized nutritional recommendations appeared to be successful in reducing weight in users with obesity [59]. Future research will focus on the inclusion of new adaptive features in health apps to achieve greater results in health outcomes and higher rates of intervention adherence.

Analyses of the Effect Stratified by Baseline Characteristics

These analyses showed that the intervention was more effective in specific groups than in the general IG. Women showed greater weight loss, BMI changes, and other anthropometric variables, whereas men did not show differences. This could be explained by the higher rate of participation by women in this study (445/650, 68.5%), following the trend of weight management studies [60] and by the lower number of men included (n=205), which may have resulted in an underpowered analysis to find differences in the male group. In addition, women are more likely to participate in weight loss interventions [61] and use health apps [62], so the sex factor has to be considered. Regarding age, people aged >50 years obtained differences in more outcomes than younger people, and they were more adherent to the self-monitoring on the app. A systematic review found that middle-aged adults are more willing to adhere to such interventions with activity trackers [63] than younger people, so it is feasible that differences are found between age groups, as in the study results. In terms of marital status, married people seem to benefit more from the intervention. However, the unequal size of the groups could explain why single people, who obtained higher median adherence days than married people, did not show positive results. The influence of sociodemographic factors on the digital intervention effect, or adherence, has to be explored in-depth, but a study among users of eHealth approaches suggested that married people, among other characteristics, generally used more mHealth apps [64], which might lead to more positive results. Although more research is needed to determine which personal factors could influence mHealth effectiveness, these analyses highlight the need to develop more tailored interventions, adapting them according to certain characteristics of the user to enhance the effect of these digital approaches.

Finally, the average dropout rate was higher than expected (207/650, 31.8%) but balanced between study groups (IG: 102/318, 32.1% vs CG: 105/332, 31.6%). As a group, the participants who dropped out of the study were younger, with greater mean weight, BMI >30 kg/m^2, and a higher proportion of smokers, in line with the dropout predictors found by a systematic review [65]. The participant attrition rate from digital interventions often exceeds 20% [66], and it is common to find large dropouts in weight management mobile phone apps [67] and multicomponent interventions [51]. Potential explanations could be participants' higher weight loss expectations [65] and experiencing difficulty in maintaining self-monitoring.

Strengths and Limitations

This study had several strengths. The study included a large sample with a multisite design with a wide range of ages and educational background, which offers robustness to the results obtained. Both the intervention and the statistical analysis were conducted by blinding researchers to the assignment groups. The 3- and 12-month visits allowed evaluation of the short-term and long-term effects of this type of technology on weight loss in the absence of additional face-to-face intervention. Moreover, the adherence rate to the self-monitoring diet on the app was acceptable (median percentage of days 71.7%).

In addition, some limitations of this study should be noted. Although participants were instructed not to use any other mHealth tool that could interfere with the study, there are no guarantees to ensure this occurred. The data collected from the smart band did not allow for the assessment of daily use and adherence to this device. In addition, the nature of the intervention precludes blinding of the participants, although recent findings suggest that blinding is less important than often believed [68]. The exposure time to the intervention (3 months) might not be sufficient to identify more positive results in

changing lifestyles and weight loss. The number of participants in each baseline characteristics (sex, age, and marital status) group could be insufficient to show more relevant effects in the stratified analysis. Finally, the dropout rate of 31.8% (207/650) may have biased the final sample study composition and underpowered the study with regard to detecting a significant effect in the results between groups. However, random allocation and a balanced dropout between arms show that the group characteristics differ little from the initial sample, making the comparison between groups possible [69].

Conclusions

The low-intensity intervention of the Evident 3 study showed in the IG, benefits on weight loss, some body composition variables, and time spent in LPA compared with the CG at 3 months, but once the devices were collected, the downward trend was not maintained at the 12-month follow-up. No differences in nutritional outcomes were observed between the groups. Analyses stratified by baseline characteristics revealed that the intervention was more effective in women, people aged ≥50 years, and married participants. Further research is needed to determine the optimum intervention period to achieve greater results, as well as the inclusion of more tailored strategies in health apps and weight management interventions that improve intervention adherence and retention rates.

Acknowledgments

This study was funded by the Spanish Ministry of Science and Innovation, Instituto de Salud Carlos III and cofunded by the European Union (ERDF/ESF, "Investing in your future"; RD16/0007/0003, RD16/0007/0005, RD16/0007/0008, and RD16/0007/0009; and PI16/00101, PI16/00952, PI16/00765, PI16/00659, PI16/00421, PI16/00170, and FI17/00040; REDIAPP). Gerencia Regional de Salud de Castilla y Leon (GRS 1277/B/16 and GRS 1580/B/17) also collaborated with the funding of this study. They played no role in the study design, data analysis, reporting results, or the decision to submit the manuscript for publication.

The Evident 3 Investigators Group comprised contributors from the following centers.

Unidad de Investigación de Atención Primaria de Salamanca (APISAL): Luis García-Ortiz (principal investigator), José I Recio-Rodríguez, Cristina Lugones-Sánchez, Manuel A Gómez-Marcos, Emiliano Rodríguez-Sánchez, Olaya Tamayo-Morales, Rosario Alonso-Domínguez, Natalia Sánchez-Aguadero, Susana González-Sánchez, Ángela de Cabo-Laso, Carmela Rodríguez-Martín, Carmen Castaño-Sánchez, Benigna Sánchez-Salgado, Jesus González-Sánchez, María C Patino-Alonso, José A Maderuelo-Fernández, Leticia Gómez-Sánchez, and Inés Llamas-Ramos.

Centro de Salud Torreramona de Zaragoza (Health Service of Aragón): Natividad González-Viejo, José Félix Magdalena-Belio, Luis Otegui-Ilarduya, Francisco J Rubio-Galán, Cristina I Sauras-Yera, Amor Melguizo-Bejar, María J Gil-Train, Marta Iribarne-Ferrer, Olga Magdalena-González, Miguel A Lafuente-Ripollés, M Mar Martínez, and Pilar Jiménez-Marcén.

Centro de Salud Cuenca I (Health Service of Castilla-La Mancha): Fernando Salcedo-Aguilar, Fructuoso Muelas-Herraiz, María A Molina-Morate, Amparo Pérez-Parra, Fernando Madero, Ángel García-Imbroda, José M Izquierdo, and María L Monterde.

Universidad de Castilla-La Mancha (University of Castilla-La Mancha): Vicente Rodríguez-Vizcaino, Alba, Soriano-Cano, Diana Patricia Pozuelo-Carrascosa, Esther Gálvez-Adalia, Alicia del Saz-Lara, and Ana Díez-Fernandez.

Centro de Salud Sta Ponça de Palma de Mallorca (Health Service of Balear Islands): José I Ramírez-Manent, José L Ferrer-Perelló, José E Romero-Palmer, Manuel Sarmiento-Cruz, Guillermo Artigues, Jitka Mudrychova, María Albaladejo-Blanco, Margarita I Moyá-Seguí, Cristina Vidal-Ribas, Patricia Lorente-Montalvo, Isabel Torrens-Darder, María M Torrens-Darder, and Lucía Pascual Calleja.

Centro de Salud San Pablo de Valladolid (Health Service of Castilla y León): María J Álvarez-Miguel, María D de Arriba-Gómez, María A Rodríguez-Fernández, Isabel Arranz-Hernando, Silvia Ramos-de la Torre, Amparo Arqueaga-Luengo, María E Moreno-Moreno, Agustina Marcos-García, Nora Manrique-Vinagre, Nieves Palomo-Blázquez, José L Montalvillo-Montalvillo, María E Fernández-Rodríguez, Alejandro González-Moro, Marta, Santiago-Pastor, María I Pérez-Concejo, and Aurora Rubio-Fernández.

Centro de Salud Casa del Barco de Valladolid (Health Service of Castilla y León): Amparo Gómez-Arranz, Carmen Fernández-Alonso, Daniel Rodríguez-Dominguez, Irene Repiso-Gento, Aventina de la Cal-de la Fuente, Rosa, Aragón-García, Miguel A Díez-Garcia, Elisa Ibañes-Jalón, Ines Castrillo-Sanz, Ana M Corcho-Castaño, Esther Jiménez-López, Daniel Correa-González, Lucía Barruso-Villafaina, Isabel Peña-García, Dolores Escudero-Terrón, Pilar Mena-Martín, Rosario Fraile-Gómez, Alberto Alonso-Gómez, Pilar Urueña, Francisca Martínez-Bermejo, Concepción Hernández-San José, Manuela Nuñez-Gómez, Patricia Sanz-Capdepont, Ana I Pazos-Revuelta, Sofía Pérez-Niño, and María E Junquera-del Pozo.

The CGB Computer Company in Salamanca, Spain, contributed to the technical development of the EVIDENT 3 app.

Authors' Contributions

LGO, MAGM, and JIRR contributed to the conception and design of the study. LGO had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of data analysis. CLS, LGO, and MAGM contributed to the drafting of the paper, and CLS had the primary responsibility for the final content. LGO and MAGM contributed as senior authors to the manuscript. LGO, ERS, and MAGM contributed to the analysis and interpretation of the data. CLS, MAGM, ERS, JIRR, and LGO contributed to the critical review of the paper for important intellectual content. CLS, JIRR, CAC, IRG, EGA, MASC, and JIRM were responsible for the collection and assembly of data. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Baseline characteristics comparison between participants who completed the study and those who dropped out.
[PDF File (Adobe PDF File), 91 KB - jmir_v24i2e30416_app1.pdf]

Multimedia Appendix 2
Adherence to the smartphone app (number of days with a record in the app).
[PDF File (Adobe PDF File), 62 KB - jmir_v24i2e30416_app2.pdf]

Multimedia Appendix 3
Adherence to self-monitoring on the app analyzed by the median percentage out of the 90 days of the intervention grouped by sex, marital status and age.
[PDF File (Adobe PDF File), 68 KB - jmir_v24i2e30416_app3.pdf]

Multimedia Appendix 4
Correlations between International Physical Activity Questionnaire-Short Form and the Smart band at the 3-month visit in the intervention group.
[PDF File (Adobe PDF File), 96 KB - jmir_v24i2e30416_app4.pdf]

Multimedia Appendix 5
Analysis of the mobile health intervention effect on weight and body composition variables grouped by baseline characteristics.
[PDF File (Adobe PDF File), 117 KB - jmir_v24i2e30416_app5.pdf]

Multimedia Appendix 6
CONSORT eHEALTH Checklist (V 1.6.1).
[PDF File (Adobe PDF File), 335 KB - jmir_v24i2e30416_app6.pdf]

References

1. Xu H, Cupples LA, Stokes A, Liu C. Association of obesity with mortality over 24 years of weight history: findings from the Framingham heart study. JAMA Netw Open 2018 Nov 02;1(7):e184587 [FREE Full text] [doi: 10.1001/jamanetworkopen.2018.4587] [Medline: 30646366]
2. Martinson BC, O'Connor PJ, Pronk NP. Physical inactivity and short-term all-cause mortality in adults with chronic disease. Arch Intern Med 2001 May 14;161(9):1173-1180. [doi: 10.1001/archinte.161.9.1173] [Medline: 11343440]
3. Stenholm S, Head J, Kivimäki M, Kawachi I, Aalto V, Zins M, et al. Smoking, physical inactivity and obesity as predictors of healthy and disease-free life expectancy between ages 50 and 75: a multicohort study. Int J Epidemiol 2016 Aug;45(4):1260-1270 [FREE Full text] [doi: 10.1093/ije/dyw126] [Medline: 27488415]
4. Lavie CJ, McAuley PA, Church TS, Milani RV, Blair SN. Obesity and cardiovascular diseases: implications regarding fitness, fatness, and severity in the obesity paradox. J Am Coll Cardiol 2014 Apr 15;63(14):1345-1354 [FREE Full text] [doi: 10.1016/j.jacc.2014.01.022] [Medline: 24530666]
5. Withrow D, Alter D. The economic burden of obesity worldwide: a systematic review of the direct costs of obesity. Obes Rev 2011 Feb;12(2):131-141. [doi: 10.1111/j.1467-789X.2009.00712.x] [Medline: 20122135]
6. Acosta A, Streett S, Kroh MD, Cheskin LJ, Saunders KH, Kurian M, et al. White Paper AGA: POWER - Practice guide on obesity and weight management, education, and resources. Clin Gastroenterol Hepatol 2017 May;15(5):631-649. [doi: 10.1016/j.cgh.2016.10.023] [Medline: 28242319]
7. Baillot A, Romain A, Boisvert-Vigneault K, Audet M, Baillargeon J, Dionne I, et al. Effects of lifestyle interventions that include a physical activity component in class II and III obese individuals: a systematic review and meta-analysis. PLoS One 2015 Apr 1;10(4):e0119017 [FREE Full text] [doi: 10.1371/journal.pone.0119017] [Medline: 25830342]

8. Hassan Y, Head V, Jacob D, Bachmann M, Diu S, Ford J. Lifestyle interventions for weight loss in adults with severe obesity: a systematic review. Clin Obes 2016 Dec;6(6):395-403. [doi: 10.1111/cob.12161] [Medline: 27788558]
9. WHO Guideline: Recommendations on Digital Interventions for Health System Strengthening. Geneva: World Health Organization; 2018.
10. Agnihothri S, Cui L, Delasay M, Rajan B. The value of mHealth for managing chronic conditions. Health Care Manag Sci 2020 Jun;23(2):185-202. [doi: 10.1007/s10729-018-9458-2] [Medline: 30382448]
11. Whitehead L, Seaton P. The effectiveness of self-management mobile phone and tablet apps in long-term condition management: a systematic review. J Med Internet Res 2016 May 16;18(5):e97 [FREE Full text] [doi: 10.2196/jmir.4883] [Medline: 27185295]
12. Bhardwaj NN, Wodajo B, Gochipathala K, Paul DP, Coustasse A. Can mHealth revolutionize the way we manage adult obesity? Perspect Health Inf Manag 2017;14(Spring):1a [FREE Full text] [Medline: 28566984]
13. Wang Y, Min J, Khuri J, Xue H, Xie B, A Kaminsky L, et al. Effectiveness of mobile health interventions on diabetes and obesity treatment and management: systematic review of systematic reviews. JMIR Mhealth Uhealth 2020 Apr 28;8(4):e15400 [FREE Full text] [doi: 10.2196/15400] [Medline: 32343253]
14. Cavero-Redondo I, Martinez-Vizcaino V, Fernandez-Rodriguez R, Saz-Lara A, Pascual-Morena C, Álvarez-Bueno C. Effect of behavioral weight management interventions using lifestyle mHealth self-monitoring on weight loss: a systematic review and meta-analysis. Nutrients 2020 Jul 03;12(7):1977 [FREE Full text] [doi: 10.3390/nu12071977] [Medline: 32635174]
15. Hutchesson MJ, Rollo ME, Krukowski R, Ells L, Harvey J, Morgan PJ, et al. eHealth interventions for the prevention and treatment of overweight and obesity in adults: a systematic review with meta-analysis. Obes Rev 2015 May;16(5):376-392. [doi: 10.1111/obr.12268] [Medline: 25753009]
16. Romeo A, Edney S, Plotnikoff R, Curtis R, Ryan J, Sanders I, et al. Can smartphone apps increase physical activity? Systematic review and meta-analysis. J Med Internet Res 2019 Mar 19;21(3):e12053 [FREE Full text] [doi: 10.2196/12053] [Medline: 30888321]
17. Richardson CR, Newton TL, Abraham JJ, Sen A, Jimbo M, Swartz AM. A meta-analysis of pedometer-based walking interventions and weight loss. Ann Fam Med 2008;6(1):69-77 [FREE Full text] [doi: 10.1370/afm.761] [Medline: 18195317]
18. Chiauzzi E, Rodarte C, DasMahapatra P. Patient-centered activity monitoring in the self-management of chronic health conditions. BMC Med 2015 Apr 09;13:77 [FREE Full text] [doi: 10.1186/s12916-015-0319-2] [Medline: 25889598]
19. Fuller D, Colwell E, Low J, Orychock K, Tobin MA, Simango B, et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. JMIR Mhealth Uhealth 2020 Sep 08;8(9):e18694 [FREE Full text] [doi: 10.2196/18694] [Medline: 32897239]
20. Kirk MA, Amiri M, Pirbaglou M, Ritvo P. Wearable technology and physical activity behavior change in adults with chronic cardiometabolic disease: a systematic review and meta-analysis. Am J Health Promot 2019 Jun;33(5):778-791. [doi: 10.1177/0890117118816278] [Medline: 30586996]
21. Brickwood K, Watson G, O'Brien J, Williams AD. Consumer-based wearable activity trackers increase physical activity participation: systematic review and meta-analysis. JMIR Mhealth Uhealth 2019 Apr 12;7(4):e11819 [FREE Full text] [doi: 10.2196/11819] [Medline: 30977740]
22. Wang JB, Cadmus-Bertram LA, Natarajan L, White MM, Madanat H, Nichols JF, et al. Wearable sensor/device (Fitbit One) and SMS text-messaging prompts to increase physical activity in overweight and obese adults: a randomized controlled trial. Telemed J E Health 2015 Oct;21(10):782-792 [FREE Full text] [doi: 10.1089/tmj.2014.0176] [Medline: 26431257]
23. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 07;13(1):127 [FREE Full text] [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
24. Spring B, Pellegrini C, McFadden HG, Pfammatter AF, Stump TK, Siddique J, et al. Multicomponent mHealth intervention for large, sustained change in multiple diet and activity risk behaviors: the make better choices 2 randomized controlled trial. J Med Internet Res 2018 Jun 19;20(6):e10528 [FREE Full text] [doi: 10.2196/10528] [Medline: 29921561]
25. Duncan M, Fenton S, Brown W, Collins C, Glozier N, Kolt G, et al. Efficacy of a multi-component m-health weight-loss intervention in overweight and obese adults: a randomised controlled trial. Int J Environ Res Public Health 2020 Aug 26;17(17):6200 [FREE Full text] [doi: 10.3390/ijerph17176200] [Medline: 32859100]
26. Patel ML, Wakayama LN, Bennett GG. Self-monitoring via digital health in weight loss interventions: a systematic review among adults with overweight or obesity. Obesity (Silver Spring) 2021 Mar;29(3):478-499. [doi: 10.1002/oby.23088] [Medline: 33624440]
27. Wang Y, Xue H, Huang Y, Huang L, Zhang D. A systematic review of application and effectiveness of mHealth interventions for obesity and diabetes treatment and self-management. Adv Nutr 2017 May;8(3):449-462 [FREE Full text] [doi: 10.3945/an.116.014100] [Medline: 28507010]
28. Fawcett E, Van Velthoven MH, Meinert E. Long-term weight management using wearable technology in overweight and obese adults: systematic review. JMIR Mhealth Uhealth 2020 Mar 10;8(3):e13461 [FREE Full text] [doi: 10.2196/13461] [Medline: 32154788]

29. Jakicic JM, Davis KK, Rogers RJ, King WC, Marcus MD, Helsel D, et al. Effect of wearable technology combined with a lifestyle intervention on long-term weight loss: the IDEA randomized clinical trial. JAMA 2016 Sep 20;316(11):1161-1171 [FREE Full text] [doi: 10.1001/jama.2016.12858] [Medline: 27654602]
30. Recio-Rodríguez JI, Martín-Cantera C, González-Viejo N, Gómez-Arranz A, Arietaleanizbeascoa MS, Schmolling-Guinovart Y, EVIDENT Group. Effectiveness of a smartphone application for improving healthy lifestyles, a randomized clinical trial (EVIDENT II): study protocol. BMC Public Health 2014 Mar 15;14:254 [FREE Full text] [doi: 10.1186/1471-2458-14-254] [Medline: 24628961]
31. Lugones-Sanchez C, Sanchez-Calavera MA, Repiso-Gento I, Adalia EG, Ramirez-Manent JI, Agudo-Conde C, EVIDENT 3 Investigators. Effectiveness of an mHealth intervention combining a smartphone app and smart band on body composition in an overweight and obese population: randomized controlled trial (EVIDENT 3 study). JMIR Mhealth Uhealth 2020 Nov 26;8(11):e21771 [FREE Full text] [doi: 10.2196/21771] [Medline: 33242020]
32. Recio-Rodriguez JI, Gómez-Marcos MA, Agudo-Conde C, Ramirez I, Gonzalez-Viejo N, Gomez-Arranz A, EVIDENT 3 investigators. EVIDENT 3 Study: a randomized, controlled clinical trial to reduce inactivity and caloric intake in sedentary and overweight or obese people using a smartphone application: study protocol. Medicine (Baltimore) 2018 Jan;97(2):e9633 [FREE Full text] [doi: 10.1097/MD.0000000000009633] [Medline: 29480874]
33. Puig-Ribera A, Martín-Cantera C, Puigdomenech E, Real J, Romaguera M, Magdalena-Belio J, EVIDENT Group. Screening physical activity in family practice: validity of the Spanish version of a brief physical activity questionnaire. PLoS One 2015;10(9):e0136870 [FREE Full text] [doi: 10.1371/journal.pone.0136870] [Medline: 26379036]
34. Voils CI, Olsen MK, Gierisch JM, McVay MA, Grubber JM, Gaillard L, et al. Maintenance of weight loss after initiation of nutrition training: a randomized trial. Ann Intern Med 2017 Apr 04;166(7):463-471 [FREE Full text] [doi: 10.7326/M16-2160] [Medline: 28241185]
35. Jensen MD, Ryan DH, Apovian CM, Ard JD, Comuzzie AG, Donato KA, American College of Cardiology/American Heart Association Task Force on Practice Guidelines, Obesity Society. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society. Circulation 2014 Jun 24;129(25 Suppl 2):S102-S138 [FREE Full text] [doi: 10.1161/01.cir.0000437739.71477.ee] [Medline: 24222017]
36. Santiago Pérez M, Hervada Vidal X, Naveira Barbeito G, Silva L, Fariñas H, Vázquez E, et al. El programa epidat: usos y perspectivas. Rev Panam Salud Publica 2010 Jan;27(1):80-82 [FREE Full text] [doi: 10.1590/S1020-49892010000100012]
37. Salas-Salvadó J, Rubio MA, Barbany M, Moreno B, Grupo Colaborativo de la SEEDO. [SEEDO 2007 Consensus for the evaluation of overweight and obesity and the establishment of therapeutic intervention criteria]. Med Clin (Barc) 2007 Feb 10;128(5):184-196. [doi: 10.1016/s0025-7753(07)72531-9] [Medline: 17298782]
38. Krakauer NY, Krakauer JC. A new body shape index predicts mortality hazard independently of body mass index. PLoS One 2012;7(7):e39504 [FREE Full text] [doi: 10.1371/journal.pone.0039504] [Medline: 22815707]
39. Thomas D, Bredlau C, Bosy-Westphal A, Mueller M, Shen W, Gallagher D, et al. Relationships between body roundness with body fat and visceral adipose tissue emerging from a new geometrical model. Obesity (Silver Spring) 2013 Nov;21(11):2264-2271 [FREE Full text] [doi: 10.1002/oby.20408] [Medline: 23519954]
40. Román Viñas B, Ribas Barba L, Ngo J, Serra Majem L. [Validity of the international physical activity questionnaire in the Catalan population (Spain)]. Gac Sanit 2013;27(3):254-257 [FREE Full text] [doi: 10.1016/j.gaceta.2012.05.013] [Medline: 23103093]
41. Jetté M, Sidney K, Blümchen G. Metabolic equivalents (METS) in exercise testing, exercise prescription, and evaluation of functional capacity. Clin Cardiol 1990 Aug;13(8):555-565 [FREE Full text] [doi: 10.1002/clc.4960130809] [Medline: 2204507]
42. Fernández-Ballart JD, Piñol JL, Zazpe I, Corella D, Carrasco P, Toledo E, et al. Relative validity of a semi-quantitative food-frequency questionnaire in an elderly Mediterranean population of Spain. Br J Nutr 2010 Jun;103(12):1808-1816. [doi: 10.1017/S0007114509993837] [Medline: 20102675]
43. Schröder H, Fitó M, Estruch R, Martínez-González MA, Corella D, Salas-Salvadó J, et al. A short screener is valid for assessing Mediterranean diet adherence among older Spanish men and women. J Nutr 2011 Jun;141(6):1140-1145. [doi: 10.3945/jn.110.135566] [Medline: 21508208]
44. ESH/ESC Task Force for the Management of Arterial Hypertension. 2013 Practice guidelines for the management of arterial hypertension of the European Society of Hypertension (ESH) and the European Society of Cardiology (ESC): ESH/ESC task force for the management of arterial hypertension. J Hypertens 2013 Oct;31(10):1925-1938. [doi: 10.1097/HJH.0b013e328364ca4c] [Medline: 24107724]
45. WHO MONICA Project Principal Investigators. The World Health Organization MONICA Project (monitoring trends and determinants in cardiovascular disease): a major international collaboration. J Clin Epidemiol 1988;41(2):105-114. [doi: 10.1016/0895-4356(88)90084-4] [Medline: 3335877]
46. Raidl M, Rich M, Lanting R, Safaii S. The healthy diabetes plate. J Am Dietetic Assoc 2010 Sep;110(9):A15. [doi: 10.1016/j.jada.2010.06.050]
47. BEDCA. Base de datos Española de composición de alimentos. URL: https://bedca.net/ [accessed 2022-01-11]

48. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
49. Greaves CJ, Sheppard KE, Abraham C, Hardeman W, Roden M, Evans PH, IMAGE Study Group. Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. BMC Public Health 2011 Feb 18;11:119 [FREE Full text] [doi: 10.1186/1471-2458-11-119] [Medline: 21333011]
50. Thomas J, Raynor H, Bond D, Luke A, Cardoso C, Foster G, et al. Weight loss in Weight Watchers Online with and without an activity tracking device compared to control: a randomized trial. Obesity (Silver Spring) 2017 Jun;25(6):1014-1021. [doi: 10.1002/oby.21846] [Medline: 28437597]
51. Kurscheid T, Redaélli M, Heinen A, Hahmann P, Behle K, Froböse I. [App-controlled feedback devices can support sustainability of weight loss. Multicentre QUANT-study shows additional weight loss and gain of QoL via multiple feedback-devices in OPTIFAST®52-program]. Z Psychosom Med Psychother 2019 Sep;65(3):224-238. [doi: 10.13109/zptm.2019.65.3.224] [Medline: 31476994]
52. Painter SL, Ahmed R, Hill JO, Kushner RF, Lindquist R, Brunning S, et al. What matters in weight loss? An in-depth analysis of self-monitoring. J Med Internet Res 2017 May 12;19(5):e160 [FREE Full text] [doi: 10.2196/jmir.7457] [Medline: 28500022]
53. Harvey J, Krukowski R, Priest J, West D. Log often, lose more: electronic dietary self-monitoring for weight loss. Obesity (Silver Spring) 2019 Mar;27(3):380-384 [FREE Full text] [doi: 10.1002/oby.22382] [Medline: 30801989]
54. Waters L, George AS, Chey T, Bauman A. Weight change in control group participants in behavioural weight loss interventions: a systematic review and meta-regression study. BMC Med Res Methodol 2012 Aug 08;12(1):120 [FREE Full text] [doi: 10.1186/1471-2288-12-120] [Medline: 22873682]
55. Füzéki E, Engeroff T, Banzer W. Health benefits of light-intensity physical activity: a systematic review of accelerometer data of the national health and nutrition examination survey (NHANES). Sports Med 2017 Sep;47(9):1769-1793. [doi: 10.1007/s40279-017-0724-0] [Medline: 28393328]
56. Swindell N, Rees P, Fogelholm M, Drummen M, MacDonald I, Martinez JA, et al. Compositional analysis of the associations between 24-h movement behaviours and cardio-metabolic risk factors in overweight and obese adults with pre-diabetes from the PREVIEW study: cross-sectional baseline analysis. Int J Behav Nutr Phys Act 2020 Mar 04;17(1):29 [FREE Full text] [doi: 10.1186/s12966-020-00936-5] [Medline: 32131847]
57. Cadmus-Bertram LA, Marcus BH, Patterson RE, Parker BA, Morey BL. Randomized trial of a Fitbit-based physical activity intervention for women. Am J Prev Med 2015 Sep;49(3):414-418 [FREE Full text] [doi: 10.1016/j.amepre.2015.01.020] [Medline: 26071863]
58. Solbrig L, Jones R, Kavanagh D, May J, Parkin T, Andrade J. People trying to lose weight dislike calorie counting apps and want motivational support to help them achieve their goals. Internet Interv 2017 Mar;7:23-31 [FREE Full text] [doi: 10.1016/j.invent.2016.12.003] [Medline: 28286739]
59. Hu EA, Nguyen V, Langheier J, Shurney D. Weight reduction through a digital nutrition and food purchasing platform among users with obesity: longitudinal study. J Med Internet Res 2020 Sep 02;22(9):e19634 [FREE Full text] [doi: 10.2196/19634] [Medline: 32792332]
60. Pagoto S, Schneider K, Oleski J, Luciani J, Bodenlos J, Whited M. Male inclusion in randomized controlled trials of lifestyle weight loss interventions. Obesity (Silver Spring) 2012 Jun;20(6):1234-1239 [FREE Full text] [doi: 10.1038/oby.2011.140] [Medline: 21633403]
61. French SA, Jeffery RW, Wing RR. Sex differences among participants in a weight-control program. Addict Behav 1994;19(2):147-158. [doi: 10.1016/0306-4603(94)90039-6] [Medline: 8036962]
62. Guertler D, Vandelanotte C, Kirwan M, Duncan MJ. Engagement and nonusage attrition with a free physical activity promotion program: the case of 10,000 steps Australia. J Med Internet Res 2015 Jul 15;17(7):e176 [FREE Full text] [doi: 10.2196/jmir.4339] [Medline: 26180040]
63. Cheatham SW, Stull KR, Fantigrassi M, Motel I. The efficacy of wearable activity tracking technology as part of a weight loss program: a systematic review. J Sports Med Phys Fitness 2018 Apr;58(4):534-548. [doi: 10.23736/S0022-4707.17.07437-0] [Medline: 28488834]
64. Elavsky S, Smahel D, Machackova H. Who are mobile app users from healthy lifestyle websites? Analysis of patterns of app use and user characteristics. Transl Behav Med 2017 Dec;7(4):891-901 [FREE Full text] [doi: 10.1007/s13142-017-0525-x] [Medline: 28929368]
65. Moroshko I, Brennan L, O'Brien P. Predictors of dropout in weight loss interventions: a systematic review of the literature. Obes Rev 2011 Nov;12(11):912-934. [doi: 10.1111/j.1467-789X.2011.00915.x] [Medline: 21815990]
66. Kozak AT, Buscemi J, Hawkins MA, Wang ML, Breland JY, Ross KM, et al. Technology-based interventions for weight management: current randomized controlled trial evidence and future directions. J Behav Med 2017 Feb;40(1):99-111 [FREE Full text] [doi: 10.1007/s10865-016-9805-z] [Medline: 27783259]
67. Laing BY, Mangione CM, Tseng C, Leng M, Vaisberg E, Mahida M, et al. Effectiveness of a smartphone application for weight loss compared with usual care in overweight primary care patients: a randomized, controlled trial. Ann Intern Med 2014 Nov 18;161(10 Suppl):S5-12 [FREE Full text] [doi: 10.7326/M13-3005] [Medline: 25402403]

68. Moustgaard H, Clayton GL, Jones HE, Boutron I, Jørgensen L, Laursen DR, et al. Impact of blinding on estimated treatment effects in randomised clinical trials: meta-epidemiological study. BMJ 2020 Jan 21;368:l6802 [FREE Full text] [doi: 10.1136/bmj.l6802] [Medline: 31964641]
69. Dumville JC, Torgerson DJ, Hewitt CE. Reporting attrition in randomised controlled trials. BMJ 2006 Apr 22;332(7547):969-971 [FREE Full text] [doi: 10.1136/bmj.332.7547.969] [Medline: 16627519]

Abbreviations

APISAL: Primary Care Research Unit of Salamanca
CG: control group
IBSAL: Biomedical Research Institute of Salamanca
IG: intervention group
IPAQ-SF: International Physical Activity Questionnaire–Short Form
LPA: light physical activity
MET: metabolic equivalent
mHealth: mobile health
REDIAPP: Network for Preventive Activity and Health Promotion

Edited by A Mavragani; submitted 13.05.21; peer-reviewed by M Patel, J Alvarez Pitti; comments to author 11.08.21; revised version received 18.10.21; accepted 30.11.21; published 01.02.22.

Please cite as:
Lugones-Sanchez C, Recio-Rodriguez JI, Agudo-Conde C, Repiso-Gento I, G Adalia E, Ramirez-Manent JI, Sanchez-Calavera MA, Rodriguez-Sanchez E, Gomez-Marcos MA, Garcia-Ortiz L, EVIDENT 3 Investigators
Long-term Effectiveness of a Smartphone App Combined With a Smart Band on Weight Loss, Physical Activity, and Caloric Intake in a Population With Overweight and Obesity (Evident 3 Study): Randomized Controlled Trial
J Med Internet Res 2022;24(2):e30416
URL: https://www.jmir.org/2022/2/e30416
doi:10.2196/30416
PMID:35103609

Original Paper

Blood Pressure Monitoring as a Digital Health Tool for Improving Diabetes Clinical Outcomes: Retrospective Real-world Study

Yifat Fundoiano-Hershcovitz[1], PhD; Dror Bacher[1], BSc, MBA; Marilyn D Ritholz[2], PhD; David L Horwitz[3], MD, PhD; Omar Manejwala[1], MD; Pavel Goldstein[4], PhD

[1]DarioHealth, Caesarea, Israel

[2]Joslin Diabetes Center, Harvard Medical School, Boston, MA, United States

[3]DLH Biomedical Consulting, Las Vegas, NV, United States

[4]School of Public Health, University of Haifa, Haifa, Israel

Corresponding Author:
Yifat Fundoiano-Hershcovitz, PhD
DarioHealth
Hatochen, 8
Caesarea, 3088900
Israel
Phone: 972 525296979
Email: yifat@mydario.com

Abstract

Background: Remote data capture for blood glucose (BG) or blood pressure (BP) monitoring and the use of a supportive digital app are becoming the model in diabetes and hypertension chronic care. One of the goals in chronic condition management is to increase awareness and generate behavioral change in order to improve outcomes in diabetes and related comorbidities, such as hypertension. In addition, there is a lack of understanding of the association between BG and BP levels when using digital health tools.

Objective: By applying a rigorous study framework to digital health data, this study investigated the relationship between BP monitoring and BG and BP levels, as well as a lagged association between BP and BG. We hypothesized that during the first 6 months of BP monitoring, BG and BP levels would decrease. Finally, we suggested a positive association between BP levels and the following month's BG levels.

Methods: In this retrospective, real-world case-control study, we extracted the data of 269 people with type 2 diabetes (T2D) who tracked their BG levels using the Dario digital platform for a chronic condition. We analyzed the digital data of the users who, in addition to BG, monitored their BP using the same app (BP-monitoring [BPM] group, n=137) 6 months before and after starting their BP monitoring. Propensity score matching established a control group, no blood pressure monitoring (NBPM, n=132), matched on demographic and baseline clinical measures to the BPM group. A piecewise mixed model was used for analyzing the time trajectories of BG, BP, and their lagged association.

Results: Analysis revealed a significant difference in BG time trajectories associated with BP monitoring in BPM and NBPM groups (t=–2.12, P=.03). The BPM group demonstrated BG reduction improvement in the monthly average BG levels during the first 6 months (t=–3.57, P<.001), while BG did not change for the NBPM group (t=0.39, P=.70). Both groups showed similarly stable BG time trajectories (B=0.98, t=1.16, P=.25) before starting the use of the BP-monitoring system. In addition, the BPM group showed a significant reduction in systolic (t=–6.42, P<.001) and diastolic (t=–4.80, P<.001) BP during the first 6 months of BP monitoring. Finally, BG levels were positively associated with systolic (B=0.24, t=2.77, P=.001) and diastolic (B=0.30, t=2.41, P=.02) BP.

Conclusions: The results of this study shed light on the association between BG and BP levels and on the role of BP self-monitoring in diabetes management. Our findings also underscore the need and provide a basis for a comprehensive approach to understanding the mechanism of BP regulation associated with BG.

(J Med Internet Res 2022;24(2):e32923) doi:10.2196/32923

KEYWORDS

blood glucose; blood pressure; monitoring; digital therapeutic; diabetes; hypertension; app; model; chronic disease; health data

Introduction

Major goals for the treatment of diabetes are to prevent or delay complications and optimize the quality of life [1]. People with type 2 diabetes (T2D) face challenging self-management regimens to improve glycemia and decrease morbidity and mortality, while often dealing with high costs of care [2]. Hypertension is the most common chronic illness in the United States, and the standard model of office-based care delivery continues to yield suboptimal outcomes, with approximately 50% of affected patients not achieving blood pressure (BP) control [3]. Elevated BP values are a common finding in people with T2D. In fact, hypertension is reported in over two-thirds of patients with T2D [4], and its development coincides with the development of hyperglycemia [5]. Furthermore, a large proportion of persons with diabetes exhibit poorly controlled hypertension, which may reflect not only delayed recognition of the presence of hypertension, clinical inaction, and poor adherence to the prescribed regimen but also uncertainty regarding the treatment targets and pathogenic correlation [6].

The pathogenic relationship between T2D and hypertension is assumed to be bidirectional [7]. Elevated BP levels are supposed to reflect at least partially the impact of the underlying insulin resistance on the vasculature and kidneys, while there is clinical evidence suggesting that disturbances in carbohydrate metabolism are more common in individuals with hypertension [4]. In persons with diabetes, hypertension confers an enhanced risk of cardiovascular disorder, having a similar risk for people with hypertension but without diabetes [5]. Hypertension and diabetes are major risk factors for cardiac diseases, stroke, and kidney disorders [8-10]; however, hypertension is the leading cardiovascular disease–related cause of morbidity and mortality among persons with T2D [11,12].

Previous studies have shown the beneficial effect of BP-lowering treatment on end-stage renal disorders in T2D [13]. Moreover, a significant improvement was demonstrated in all diabetes-related outcomes resulting from long-term tight BP control in patients with T2D and hypertension [14,15].

Among normotensive individuals, T2D at baseline was shown as a significant predictor of incident hypertension (independent of age, body mass index [BMI], and family history of diabetes) [6]. Furthermore, most of the care for the patient with hypertension typically relies on the primary care physician, whose time for face-to-face patient care has become progressively limited [3]. Approximately half of the persons treated for T2D do not have adequate BP or glycemic control [16]. It is clear that management of a chronic condition requires a change in strategy to meet the real-time needs of the population.

Ideal management of chronic conditions, such as T2D and hypertension, often includes monitoring lifestyle changes and pharmacological interventions to improve metabolic health [17]. Home BP measurement has been recommended by many hypertension guidelines and addresses several limitations of traditional office-based care, including reducing misclassification because of white-coat or masked hypertension and an ability to take a more suitable action and a course of corrective therapy [18].

Self-monitoring has been shown in previous studies as 1 of the key elements for successful chronic condition management. One of the examples of successful chronic condition management was shown with self-monitoring behaviors involving weight measurements, which demonstrated self-monitoring as a significant predictor of weight loss during 6 months [19]. Of note, there is evidence showing that patients' willingness to self-monitor is associated with disease controllability, and persons with diabetes, asthma, and hypertension are most willing to self-monitor [20]. Home glucose self-monitoring has been associated with improved glycemic control and reduced long-term complications [21]. Current meta-analysis supports the claim that self-monitoring can significantly reduce BP [22,23].

Treatment optimization through digital health could enhance users' alertness to their health conditions through real-time monitoring, leading to effective treatments. Timely communication and feedback also can play a key role in efforts toward achieving hypertension and diabetes control. Technology-driven solutions can help persons with diabetes build awareness of their daily health-related behaviors and promote increased engagement with those behaviors [24-26]. Communication of test results also has been shown to be highly desired among people with hypertension [27], and lifestyle-focused educational messages providing advice, motivational reminders, and support also were shown to be effective in improving hypertension and other chronic conditions [28]. Real-time digital communication can additionally include progress reports focusing on achieving BP and blood glucose (BG) goals and displaying insights as well as guidelines promoting lifestyle change. Using a mobile app for self-management purposes could make it easier for people with chronic conditions to obtain insight into and control of their BG and BP levels.

On the one hand, mobile apps have been shown to improve diabetes outcomes via education and support for adhering to evidence-based recommendations [29-32]. On the other hand, there are mobile apps focused solely on hypertension management. These are designed primarily for health management functions and have proved an effective solution in improving medication adherence and systolic and diastolic BP levels [33-35]. Although mobile apps have the potential to be beneficial for people with hypertension or diabetes, little is known about their efficacy targeting several chronic conditions at 1 time or monitoring both BP and BG levels on 1 mobile app [36]. Further research is necessary for investigating the effectiveness of mobile health for hypertension self-management over time [35]. Moreover, the current literature is missing rigorous real-life studies to test the role of BP self-monitoring and diabetes management and to better understand the direct association between BP monitoring and glycemic outcomes. Further, in glucose-lowering trials, it is not clear whether different ongoing BP levels are associated with different BG-lowering outcome effects [37,38].

This study used a retrospective analysis of a home-use diabetes BG meter and BP-monitoring system with full data capture using a supportive mobile app among people with T2D and poorly controlled BP levels. This study can illuminate the dynamic of the relative contribution of BP self-monitoring to successful diabetes and hypertension self-management using real-life data. We analyzed users' data for 6 months before and 6 months after using the BP-monitoring system and compared them with a matched control group that never used the BP-monitoring system (NBPM). We hypothesized that BP monitoring will be associated with reduction in systolic and diastolic BP, as well as with BG levels. We additionally hypothesized a linkage between BP levels and following BG levels.

Methods

Platform

This study used the Dario digital therapeutics solution (DarioHealth) for chronic conditions to support self-management of diabetes and high BP levels. The Dario BP-monitoring system combines an innovative meter with a phone app that is available for both Android and iOS devices. The glucose meter consists of a small, pocket-size holder for strips, a lancet, and the meter. The BG meter is removed from the holder and plugged directly into a smart mobile device, effectively converting the smart mobile device into a display screen for the meter. The BP-monitoring system measures the systolic and diastolic BP and pulse rate by using a noninvasive technique in which an inflatable cuff is wrapped on the upper arm. The BP-monitoring system provides Bluetooth transmission. The BP cuff is paired with the mobile app, and the data are transmitted to the smart mobile device via Bluetooth. The BP reading is displayed on the mobile app screen.

First, connecting the BG meter directly and pairing the BP cuff to the phone ensures 100% data capture during glucose readings. Second, users open the mobile app with their BG or BP measurement. The measurements are taken independently. This makes contextually tagging a measurement easy at the time of taking the measurement. More specifically, the measurement is shown on the mobile phone in a decision support system view. After the measurement is shown, the user is transferred to a data entry screen where additional information (measurement time [fasting/premeal/postmeal/bedtime]; carbohydrate intake (g); meal, mood, and location settings; and physical activity [kcal]) can be added to the BG measurement. All information is stored in the users' logbook in the app, attached to the specific BG or BP reading. Data are uploaded to the cloud for backup and further analysis. Mobile app functions include interface design elements as well as specific educational content, wording, or digital interventions that affect the users' choices in the digital environment that provides personal health information and prompt feedback.

Measures

The outcome metrics were the monthly average BP level (systolic and diastolic BP), defined as the mean of all the user's BP measurements taken over a 30-day interval, and the monthly average BG level, defined as the mean of all the user's BG measurements taken over a 30-day interval.

The relationships of interest potentially could be investigated on different scales emphasizing daily, weekly, or monthly fluctuations. In this real-world data analysis, the timescale was designed to reflect the monthly aggregated interval change over a 6-month period because of the difficulty in tracking daily changes in digital monitoring.

The mobile app collected the following medical and sociodemographic information (by self-report) for each user: gender, age, BMI, physical activity, stress level, and comorbidities, such as high lipids, chronic kidney disease, cardiovascular disease, sleep disorder, cancer, or stress and depression. Socioeconomic status was matched by applying zip code data to census.gov. All data were transferred and stored in compliance with Health Insurance Portability and Accountability Act (HIPAA) requirements, using Amazon AWS database services. All data were anonymized before extraction for this study.

Users

In total, 269 users with T2D who used the Dario BP-monitoring system between 2019 and 2020 were included in this analysis. The sample included 172 (63.9%) men, 45 (16.7%) having comorbidities. Their average age was 62.0 (SD 11.9) years, average BMI was 31.7 (SD 6.4), and median household income was US $29,100 (SD 3150).

Study Design

The BPM group included persons with diabetes who measured their BG and BP levels (BPM group, n=137, 50.9%). Inclusion criteria were as follows for the BPM group: measured BP levels in the first and fourth months, with at least 5 BP measurements per month; the first-month average BP level was in the elevated category (systolic BP 120-129 mmHg; diastolic BP less than 80 mmHg) [18] or above; and measured BG levels in the first and fourth months after starting to use the Dario BP-monitoring system, with at least 5 BG measurements per month.

We applied a quasi-experimental case-control study design to improve the methodological rigor and validity of the findings and take advantage of the users' digital follow-up. We used the existing Dario database to extract the background population to match a control group, NBPM, that did not use BP monitoring. The BG measurement inclusion criterion for the NBPM group was the same as for the BPM group (measured BG levels in the first and fourth months, with at least 5 BG measurements per month) to find the best match for the cohort of 137 (49.1%) users who started using BP monitoring. Matching is used in the context of estimating the causal effect of a binary condition of interested in or exposed to on an outcome, while controlling for potential confounding variables or variables prognostic of the outcome [39]. The goal of matching was to produce a covariate balance, seeking for approximately equal distributions of covariates in the 2 groups, as they would be in a randomized experiment. The covariate balance results in increased robustness to the choice of model used to estimate the treatment effect. The match was based on sociodemographic and clinical parameters: gender, age, BMI,

physical activity level, stress level, and self-reported comorbidities (hypertension, high lipids, chronic kidney disease, cardiovascular disease, sleep disorder, cancer, mental condition), socioeconomic status, number of BG measurements, and average BG. We applied the nearest-neighbor propensity score matching without replacement, with the propensity score estimated using logistic regression of the treatment on the covariates [40,41]. This approach resulted in an adequate balance using the data of 132 users who did not measure BP.

No significant differences were found between BPM and NBPM conditions by age (B=–0.007, Z=–0.22, *P*=.83), gender (B=0.53, Z=0.59, *P*=.56), BMI (B=0.03, Z=0.34, *P*=.73), physical activity level (0=not active, 10=very active; B=–0.28, Z=–1.55, *P*=.12), stress level (0=not stressed, 10 =very stressed; B=0.002, Z=0.01, *P*=.99), median household income (B=–0.28, Z=–1.55, *P*=.12), number of BG measurements (B=–0.01, Z=–0.60, *P*=.55), insulin treatment (B=–0.98, Z=–0.90, *P*=.37), comorbidities (B=–0.66, Z=–0.65, *P*=.52), and digital engagement (tagging meal type, physical activity in the context of measurement [26]; B=0.02, Z=0.97, *P*=.33).

Ethical & Independent Review Services [33], a professional review board, issued the institutional review board exemption for this study (#18032-04).

Analytic Approach

A classical linear longitudinal model assumes a single-slope growth pattern for changes in an outcome variable across time. Sometimes, such a simple model does not fit the empirical data. In contrast, piecewise-based mixed-effects models allow flexibility in the modeling of trajectories across time [42]. Here, a mixed piecewise model assessed differences in the monthly average BG level in 2 segments: before and after BP-monitoring system usage. The piecewise model allowed the data to exhibit different linear trends over their different regions. This statistical approach provided an opportunity to model curvilinear changes in the monthly average BG level as a single process and to test complex effects.

Users' data were centered around the beginning of the BP measurements and 6 months before and after that point were included in the analysis. For the NBPM (control) group, that had never started BP measurements, we included users with at least 18 months of monitoring, choosing a random cutoff point and including in the analysis only the data collected during 6 months before and after the simulated cutoff point.

A piecewise-based mixed - effects model was fit to the data, modeling temporal changes of the monthly average BG level for the 2 groups (BPM vs NBPM). The piecewise cutoff point for the model was set to the beginning of BG monitoring, assuming a change in the time-related monthly average BG trajectory between the groups by the included interaction terms between time trajectories and group. The model included a person-based random intercept and random slope for the time trajectory after the piecewise cutoff.

Next, we used mixed-model analysis to access the time trajectory of systolic and diastolic BP for the initial 6 months of BP monitoring. The models included a random intercept and random slope of the time trajectory. We reported unstandardized regression weights (B), test statistics (*t*), and associated significance (*P*).

To better understand the dynamic of BP and BG association, we conducted a lagged analysis, predicting the following month's BG level based on the BP level.

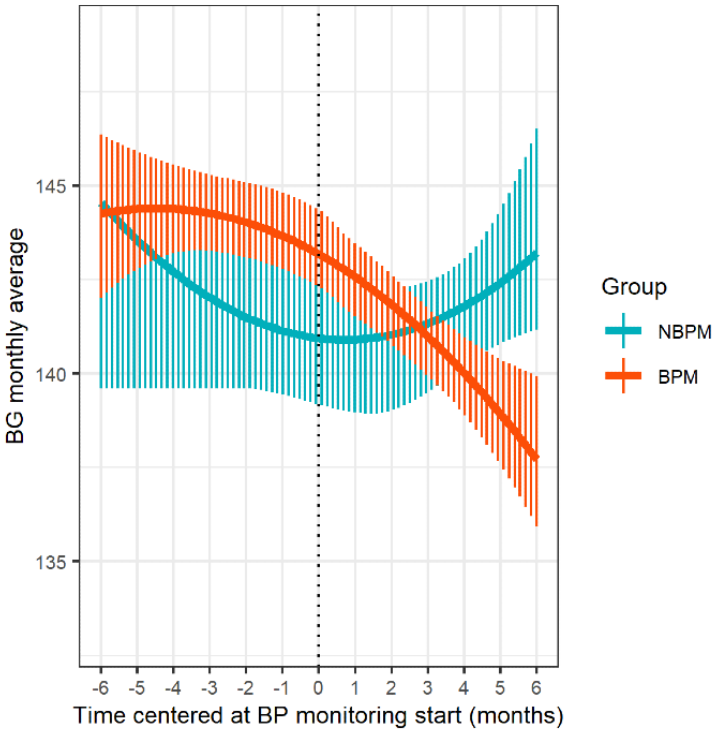
Results

BP Monitoring Is Associated With BG Levels

Piecewise mixed-model analysis revealed a significant interaction between the time after starting BP monitoring and group (B=−1.50, *t*=–2.12, *P*=.03) on BG levels. The BPM group showed a significant reduction in BG (B=−1.16, *t*=–3.57, *P*<.001), while the NBPM group did not show a significant time trend (B=0.24, *t*=0.39, *P*=.70); see Figure 1. Before BP monitoring, group difference was observed in BG time trends (B=0.98, *t*=1.16, *P*=.25), both BPM (B=−0.13, *t*=–0.43, *P*=.67) and NBPM (B=−1.16, *t*=–1.49, *P*=.14) groups showed no BG trend. Extended information is provided in Multimedia Appendix 1.

We reran the analysis including all the potential confounders into the models. In the new analysis, the pattern of the findings remained the same. Users with hypertension (B=25.27, *t*=2.57, *P*=.02) and insulin treatment (B=18.17, *t*=3.02, *P*=.003) showed increased monthly average BG levels. In addition, stress level (B=4.26, *t*=2.38, *P*=.02) and median household income (B=0.006, *t*=4.88, *P*<.001) were associated with increased monthly average BG levels. Age, gender, BMI, physical activity, alcohol consumption, and comorbidities were not related to the monthly average BG (all *P*>.09).

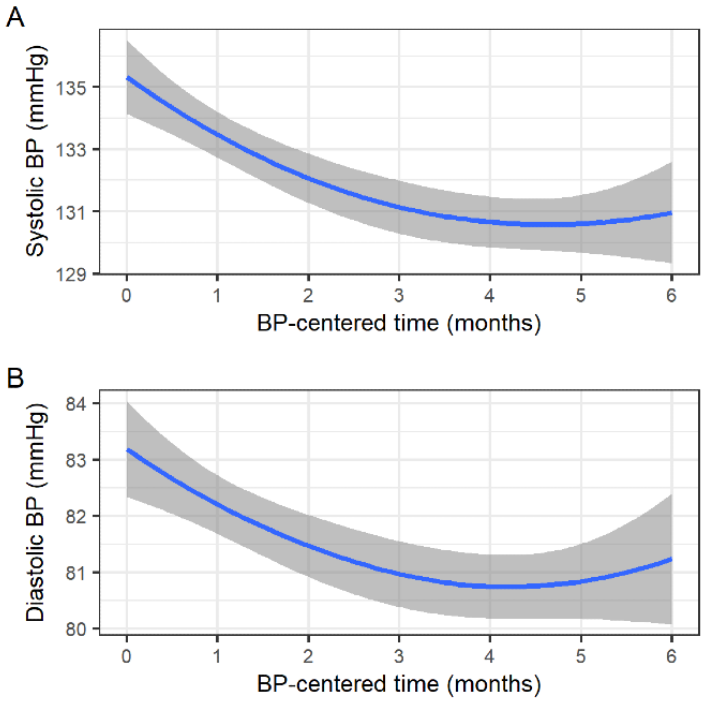
Figure 1. BG monthly average fluctuation for the BPM group and the NBPM group. Zero in the x-axis means the start of BP monitoring. Vertical lines represent a 95% CI over time. BG: blood glucose; BPM: blood pressure monitoring; NBPM: no blood pressure monitoring.



BP Monitoring and the Link to BG

We analyzed the recorded monthly averaged BP during the first 6 months of monitoring. Systolic (B=−0.82, *t*=–6.42, *P*<.001, 135.4-130.8 mmHg) and diastolic (B=−0.41, *t*=–4.80, *P*<.001, 83.3-81.7 mmHg) BP showed significant reductions during that period (Figure 2), and 37 (27%) of 137 users achieved systolic BP reduction of >10 mmHg (*P*<.001). In addition, results from the lagged analysis, predicting the following month's BG based on BP levels, showed that following month's elevated BG is associated with higher systolic (B=0.24, *t*=2.77, *P*=.001) and diastolic (B=0.30, *t*=2.41, *P*=.02) BP.

Figure 2. Monthly average BP fluctuation over the first 6 months of monitoring for (A) systolic and (B) diastolic BP. The x-axis presents the centered time, with 0 indicating the start of BP monitoring. The gray area represents a 95% CI over time. BP: blood pressure.



Discussion

Principal Results

This study showed that using a piecewise mixed-model statistical framework appears to be an appropriate base model to describe nonlinear fluctuations in BG levels comparing different user cohorts over time. In our study, the model indicated that before the BP adoption phase, both groups evidenced flat trajectories. However, after starting the use of the BP monitoring system connected to the same mobile app, the BPM group experienced a significant decrease in BG levels, while the NBPM group's BG levels remained flat. In addition, we ran a lagged analysis demonstrating that monthly systolic and diastolic BP can predict the following month's average BG. This finding suggests the hypothesis that BP reduction may serve as a mechanism of BG reduction; further studies would be needed to confirm this and to analyze the mechanism by which BP reduction has an effect on BG levels.

This real-world analysis presents data analyzing associations between BP monitoring and reduction in BG levels in people with T2D and elevated BP for those who start using the mobile app for diabetes and hypertension management. More specifically, results indicate an association between BG reduction and BP monitoring. This effect was not observed among the users who did not use BP monitoring, although both groups showed statistically equivalent BG trends before BP monitoring and had similar demographic and clinical characteristics. In addition, analysis of the monthly average BP during the first 6 months of monitoring showed a significant decrease in systolic and diastolic BP during this period. Moreover, monthly averaged BG was associated with systolic and diastolic BP levels.

Consistent with the literature, we found that the BPM group, which monitored their BP and improved their levels over time, experienced a change in their BG levels, while the NBPM group's levels remained flat over the same time. Moreover, based on the lagged analysis, we observed that higher systolic and diastolic BP is associated with elevated BG in the following month. Previous studies have shown a significant improvement in all diabetes-related outcomes resulting from long-term tight BP control in patients with T2D and hypertension [14,15]. There is substantial overlap between T2D and hypertension in etiology and illness mechanisms. Hypertension and diabetes substantially share common pathways, such as obesity, inflammation, oxidative stress, insulin resistance, and mental stress [43,44]. Patients with diabetes experience increased peripheral artery resistance caused by vascular remodeling and increased body fluid volume associated with insulin resistance–induced hyperinsulinemia and hyperglycemia. Both these mechanisms elevate systemic BP [7].

The constellation of metabolically related abnormalities, such as obesity, glucose intolerance, and hypertension, are evident in metabolic syndrome [45]. Successful management should address all the factors involved. The goal in a clinical setting is to improve the ability to identify and intervene with factors contributing to metabolic syndrome, including lifestyle modification, weight management, diet, and physical activity changes. One critical view on the individualization of target goals for patients at risk is that this has not always been well adapted for people with low education levels or other sociocultural factors that could distract from finding the time and motivation for improving their individual lifestyles [46]. Our participants showed elevated BP, BG, and BMI. Such a clinical profile resonates with previous diabetes studies [47]. Due to the nature of metabolic disease, factors such as high BP, diabetes, and obesity are shown to be the most suitable therapy areas to address via digital health. This approach may provide information suitable for the user's clinical condition and may improve self-management efficiency and the clinical course through personalized intervention [48]. Lifestyle changes, such as dietary habits and following exercise recommendations, are crucial for persons with diabetes and hypertension, even when pharmacological treatment is required [43,49]. Further, guidance on lifestyle changes must be provided for people with diabetes and hypertension.

BP management presents a possible course to control BG and improve the patient's well-being. Reducing BP is more helpful for people with than without diabetes in terms of absolute cardiovascular risk [43]. Previous studies have revealed the beneficial effect of BP treatment in people with high-risk diabetes: attaining standard glycemic control resulted in a decreased risk of cardiovascular events, including heart failure [50].

Cardiovascular disease is the most significant health threat to adults with T2D, and it is clear that efforts aimed at controlling BP and cholesterol will have much greater effects on health outcomes than those focused only on glycemic control [51]. For reducing cardiovascular risk, treatment should focus on specific goals: reductions in systolic and diastolic BP, BG, and plasma low-density lipoprotein (LDL) cholesterol [45].

Our results also revealed a synergistic effect in that the management of BG and the management of BP became evident with the start of controlling BP levels, while this association was not present in the NBPM group, which did not use the BP-monitoring system. Moreover, we demonstrated that the diabetes clinical outcome (ie, monthly averaged BG) is positively lagged with the hypertension clinical outcome (ie, systolic and diastolic BP levels).

It was previously shown how distinct features of a digital therapeutic app have the potential to deliver equitable person-centric care and how digital engagement can play a key role in enhancing a person's chronic condition self-management [16,52-54]. Previously, we demonstrated that digital engagement may improve diabetes management [26]. Importantly, in this study, the BPM and NBPM groups were not different in their digital engagement. In addition, the tracking tool can be disseminated via a simple-to-use, accessible, and low-cost device. Further, the median household distribution of the users in both groups was equal and revealed that the digital diabetes management solution is desired and affordable among lower-, middle-, and high-income levels to improve glycemic outcomes.

From a psychological perspective, it is assumed that individuals using a digital platform may develop more active roles in managing their health [55]. Previous studies have identified

psychosocial factors that could be targeted by interventions to improve diabetes self-management and treatment outcomes [51]. National survey data from the United States highlight basic knowledge gaps among many adults with diabetes and note that less than 50% know their level of glycemic control and only 63% know their BP level [56]. For people with hypertension, evidence for digital interventions has mostly come from small trials with relatively short follow-up and substantial heterogeneity of results [57]. Thus, more studies are needed with larger samples and longer periods of follow-up. Digital interventions (eg, apps, programs, or health software) have the potential to support people in self-management and to facilitate lifestyle change [57].

Our finding on the significant decrease in systolic and diastolic BP occurring within the first 6 months of using the device and mobile app suggests that even the short-term use of our digital monitoring device may be an effective means to increase users' knowledge base and self-care behavioral awareness in the context of everyday life. The mobile app provided the users with numeracy skills, including the ability to interpret and respond to hypertension numerical feedback and to focus greater attention on the self-management of more than 1 chronic condition. The intervention is designed to influence user behavior by using a person-based approach. Users were advised with personalized reminders to take BP readings, with specific messages driven by their BP levels, calculated averages, displayed BP measurements levels by color scale, and other lifestyle activities (smoking, caffeine, activity). Health behavior change theory posits that new health behaviors emerge when people gain both knowledge and self-efficacy to implement said knowledge [58-60].

The logbook screen inside the mobile app presenting measurements and data records possibly can be shared with the health care provider for further support. The digitalization of health care has the potential to save time and money and enable better physician-patient relationships and personalized treatments based on the specific characteristics of patients, especially patients with hypertension [61,62]. Important components may play a role in regulation of BP and BG and other self-measured values by health care providers.

Finally, our findings indicated that monitoring several chronic conditions may have the potential to offer a greater means for helping person with diabetes and hypertension effectively modulate their glycemia and BP than managing each of the conditions separately. We expect that the analytical approach applied in this study will be useful for examining other chronic conditions and metabolic syndrome outcomes (eg, lipid profile and weight loss). Moreover, this type of analysis may provide valuable information for optimizing patients' planning and strategies for chronic condition management.

Limitations

We noted several limitations in this study. First, as in all studies involving retrospective real-world data, groups were not randomly assigned, and treatment protocols were not prescribed. Both limitations created challenges for drawing casual effects. It is possible that users who chose to manage both diabetes and hypertension were those who were motivated to change. However, our inclusion criteria were designed to ensure that both BPM and NBPM groups showed evidence of being engaged with their diabetes management. There were no significant differences between the groups in terms of the number of BG measurements and digital engagement (behavioral tagging). This would suggest that motivation may not be the primary difference between BPM and NBPM. That said, the statistical modeling covers the pitfalls of the comparison between the 2 groups, allowing a quasi-causal inference. However, there might be variables that were not collected that may impact the group imbalance.

In this real-world data analysis, the timescale was designed to reflect the monthly interval change over the 6-month period before and the 6-month period after starting the use of the BP-monitoring system. However, the relationship of interest in this study could be potentially investigated on different scales emphasizing daily, weekly, or monthly fluctuations. Owing to the difficulty in tracking daily changes in real-world studies, most studies focus on monthly fluctuations.

Another challenge regarding our data was that available demographic data were limited. Although there were no differences between groups by age, gender, or median household income, there is always the possibility that uncontrolled demographic bias was present from other demographic factors. In addition, available medical and physical data were limited. No differences existed between groups in terms of physical activity, stress level, insulin treatment, and other comorbidities. However, there is always the possibility that an uncontrolled parameter bias was present from other medical record factors.

Conclusion

Our findings show a significant association between BP measurement and improved glycemic control. The association is consistent with the hypothesis that there is a physiologic link between BP and glucose control. An alternative explanation is that persons who measure both parameters are more likely to be involved and motivated in their health care. Focusing on BP self-monitoring and lifestyle activities may lead to better glycemic outcomes. The clinical impact was observed in users who measured their BG in the first 6-month period of BP monitoring. We also observed real-time linkage between a reduction in BP levels and a reduction in BG levels. From the behavioral science perspective, this is not surprising. Simultaneously, directing effort onto actionable areas for improvement of BP is likely to increase the thought and action needed for improvement of BG. Moreover, the process of BP self-monitoring in lowering systolic and diastolic BP levels was demonstrated. Future work should focus on investigating the mechanisms underlying the comorbidity of diabetes and hypertension and their management, identifying and applying mediation models and behavioral interventions that go beyond actionable multiple chronic conditions that drive prohealth behavioral change. Furthermore, similar studies examining the impact of gradual trajectories on other behavior changes, including health coaching, gamification, and behavioral economics, are essential. These investigations would help move the field beyond the claim of "what is the impact of the digital tools on managing chronic conditions such as diabetes or

hypertension" to a deeper understanding of how digital solutions drive clinical outcomes and how to integrate multiple digital solutions and under what clinical situations. Finally, qualitative research is needed for understanding users' real-world experiences and as a tool for sensitivity analysis and validation of complex computational models in order to enhance personalized medical approaches.

Conflicts of Interest

YFH, DB, and OM are employees of DarioHealth. MDR and DLH serve as DarioHealth scientific advisory board members, and PG has received a consulting fee to assist with analyses but otherwise has no conflict of interest.

Multimedia Appendix 1
Piecewise mixed-model analysis of BP monitoring affecting monthly BG. BG: blood glucose; BP: blood pressure.
[DOCX File , 15 KB - jmir_v24i2e32923_app1.docx]

References

1. American Diabetes Association. 4. Comprehensive medical evaluation and assessment of comorbidities. Diabetes Care 2020 Jan;43(Suppl 1):S37-S47. [doi: 10.2337/dc20-S004] [Medline: 31862747]
2. Amante DJ, Harlan DM, Lemon SC, McManus DD, Olaitan OO, Pagoto SL, et al. Evaluation of a diabetes remote monitoring program facilitated by connected glucose meters for patients with poorly controlled type 2 diabetes: randomized crossover trial. JMIR Diabetes 2021 Mar 11;6(1):e25574 [FREE Full text] [doi: 10.2196/25574] [Medline: 33704077]
3. Milani RV, Bober RM, Milani AR. Hypertension management in the digital era. Curr Opin Cardiol 2017 Jul;32(4):373-380. [doi: 10.1097/HCO.0000000000000405] [Medline: 28306675]
4. Cloutier L, Lamarre-Cliche M. Hypertension in adults with type 2 diabetes: a review of blood pressure measurement methods, targets and therapy. Can J Diabetes 2018 Apr;42(2):188-195. [doi: 10.1016/j.jcjd.2018.01.012] [Medline: 29602406]
5. Ferrannini E, Cushman WC. Diabetes and hypertension: the bad companions. Lancet 2012 Aug 11;380(9841):601-610. [doi: 10.1016/S0140-6736(12)60987-8] [Medline: 22883509]
6. Tsimihodimos V, Gonzalez-Villalpando C, Meigs JB, Ferrannini E. Hypertension and diabetes mellitus: coprediction and time trajectories. Hypertension 2018 Mar;71(3):422-428 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.117.10546] [Medline: 29335249]
7. Ohishi M. Hypertension with diabetes mellitus: physiology and pathology. Hypertens Res 2018 Jun;41(6):389-393. [doi: 10.1038/s41440-018-0034-4] [Medline: 29556093]
8. American Diabetes Association. 9. Cardiovascular disease and risk management. Diabetes Care 2017 Jan;40(Suppl 1):S75-S87. [doi: 10.2337/dc17-S012] [Medline: 27979896]
9. Benjamin EJ, Blaha MJ, Chiuve SE, Cushman M, Das SR, Deo R, American Heart Association Statistics Committee: Stroke Statistics Subcommittee. Heart disease and stroke statistics-2017 update: a report from the American Heart Association. Circulation 2017 Mar 07;135(10):e146-e603 [FREE Full text] [doi: 10.1161/CIR.0000000000000485] [Medline: 28122885]
10. Hallberg SJ, McKenzie AL, Williams PT, Bhanpuri NH, Peters AL, Campbell WW, et al. Effectiveness and safety of a novel care model for the management of type 2 diabetes at 1 year: an open-label, non-randomized, controlled study. Diabetes Ther 2018 Apr;9(2):583-612 [FREE Full text] [doi: 10.1007/s13300-018-0373-9] [Medline: 29417495]
11. Akalu Y, Belsti Y. Hypertension and its associated factors among type 2 diabetes mellitus patients at Debre Tabor General Hospital, Northwest Ethiopia. Diabetes Metab Syndr Obes 2020;13:1621-1631 [FREE Full text] [doi: 10.2147/DMSO.S254537] [Medline: 32494180]
12. Emdin CA, Rahimi K, Neal B, Callender T, Perkovic V, Patel A. Blood pressure lowering in type 2 diabetes: a systematic review and meta-analysis. JAMA 2015 Feb 10;313(6):603-615. [doi: 10.1001/jama.2014.18574] [Medline: 25668264]
13. Thomopoulos C, Parati G, Zanchetti A. Effects of blood-pressure-lowering treatment on outcome incidence in hypertension: 10; should blood pressure management differ in hypertensive patients with and without diabetes mellitus? Overview and meta-analyses of randomized trials. J Hypertens 2017 May;35(5):922-944. [doi: 10.1097/HJH.0000000000001276] [Medline: 28141660]
14. Mogensen CE. Combined high blood pressure and glucose in type 2 diabetes: double jeopardy. British trial shows clear effects of treatment, especially blood pressure reduction. BMJ 1998 Sep 12;317(7160):693-694 [FREE Full text] [doi: 10.1136/bmj.317.7160.693] [Medline: 9732334]
15. UK Prospective Diabetes Study Group. Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes: UKPDS 38. BMJ 1998 Sep 12;317(7160):703-713 [FREE Full text] [Medline: 9732337]
16. Frias J, Virdi N, Raja P, Kim Y, Savage G, Osterberg L. Effectiveness of digital medicines to improve clinical outcomes in patients with uncontrolled hypertension and type 2 diabetes: prospective, open-label, cluster-randomized pilot clinical trial. J Med Internet Res 2017 Jul 11;19(7):e246 [FREE Full text] [doi: 10.2196/jmir.7833] [Medline: 28698169]
17. Passarella P, Kiseleva TA, Valeeva FV, Gosmanov AR. Hypertension management in diabetes: 2018 update. Diabetes Spectr 2018 Aug;31(3):218-224 [FREE Full text] [doi: 10.2337/ds17-0085] [Medline: 30140137]

18. Whelton P, Carey R, Aronow W. ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association task force on clinical practice. Kidneys 2018 Feb 26;7(1):68-74. [doi: 10.22141/2307-1257.7.1.2018.122220]
19. Painter SL, Ahmed R, Hill JO, Kushner RF, Lindquist R, Brunning S, et al. What matters in weight loss? An in-depth analysis of self-monitoring. J Med Internet Res 2017 May 12;19(5):e160 [FREE Full text] [doi: 10.2196/jmir.7457] [Medline: 28500022]
20. Huygens MWJ, Swinkels ICS, de Jong JD, Heijmans MJWM, Friele RD, van Schayck OCP, et al. Self-monitoring of health data by patients with a chronic disease: does disease controllability matter? BMC Fam Pract 2017 Mar 20;18(1):40 [FREE Full text] [doi: 10.1186/s12875-017-0615-3] [Medline: 28320330]
21. Taylor J, Campbell K. Home monitoring of glucose and blood pressure. Am Fam Physician 2007 Jul 15;76(2):255-260 [FREE Full text] [Medline: 17695570]
22. Bray EP, Holder R, Mant J, McManus RJ. Does self-monitoring reduce blood pressure? Meta-analysis with meta-regression of randomized controlled trials. Ann Med 2010 Jul;42(5):371-386. [doi: 10.3109/07853890.2010.489567] [Medline: 20504241]
23. Bryant KB, Sheppard JP, Ruiz-Negrón N, Kronish IM, Fontil V, King JB, et al. Impact of self-monitoring of blood pressure on processes of hypertension care and long-term blood pressure control. J Am Heart Assoc 2020 Aug 04;9(15):e016174 [FREE Full text] [doi: 10.1161/JAHA.120.016174] [Medline: 32696695]
24. Cahn A, Akirov A, Raz I. Digital health technology and diabetes management. J Diabetes 2018 Jan;10(1):10-17. [doi: 10.1111/1753-0407.12606] [Medline: 28872765]
25. Han J, King F, Klonoff D, Drincic A, Crosby KP, Robinson T, et al. Digital Diabetes Congress 2019. J Diabetes Sci Technol 2019 Sep;13(5):979-989 [FREE Full text] [doi: 10.1177/1932296819872107] [Medline: 31466480]
26. Fundoiano-Hershcovitz Y, Hirsch A, Dar S, Feniger E, Goldstein P. Role of digital engagement in diabetes care beyond measurement: retrospective cohort study. JMIR Diabetes 2021 Feb 18;6(1):e24030 [FREE Full text] [doi: 10.2196/24030] [Medline: 33599618]
27. LaRocque JR, Davis CL, Tan TP, D'Amico FJ, Merenstein DJ. Patient preferences for receiving reports of test results. J Am Board Fam Med 2015;28(6):759-766 [FREE Full text] [doi: 10.3122/jabfm.2015.06.150030] [Medline: 26546651]
28. Chow CK, Redfern J, Hillis GS, Thakkar J, Santo K, Hackett ML, et al. Effect of lifestyle-focused text messaging on risk factor modification in patients with coronary heart disease: a randomized clinical trial. JAMA 2015;314(12):1255-1263. [doi: 10.1001/jama.2015.10945] [Medline: 26393848]
29. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
30. Offringa R, Sheng T, Parks L, Clements M, Kerr D, Greenfield M. Digital diabetes management application improves glycemic outcomes in people with type 1 and type 2 diabetes. J Diabetes Sci Technol 2018 May;12(3):701-708 [FREE Full text] [doi: 10.1177/1932296817747291] [Medline: 29277103]
31. Bollyky JB, Melton ST, Xu T, Painter SL, Knox B. The effect of a cellular-enabled glucose meter on glucose control for patients with diabetes: prospective pre-post study. JMIR Diabetes 2019 Oct 7;4(4):e14799. [doi: 10.2196/14799]
32. Osborn CY, van Ginkel JR, Rodbard D, Heyman M, Marrero DG, Huddleston B, et al. One drop | mobile: an evaluation of hemoglobin A1c improvement linked to app engagement. JMIR Diabetes 2017 Aug 24;2(2):e21 [FREE Full text] [doi: 10.2196/diabetes.8039] [Medline: 30291059]
33. Kang H, Park H. A mobile app for hypertension management based on clinical practice guidelines: development and deployment. JMIR Mhealth Uhealth 2016 Feb 02;4(1):e12 [FREE Full text] [doi: 10.2196/mhealth.4966] [Medline: 26839283]
34. Kumar N, Khunger M, Gupta A, Garg N. A content analysis of smartphone-based applications for hypertension management. J Am Soc Hypertens 2015 Feb;9(2):130-136 [FREE Full text] [doi: 10.1016/j.jash.2014.12.001] [Medline: 25660364]
35. Li R, Liang N, Bu F, Hesketh T. The effectiveness of self-management of hypertension in adults using mobile health: systematic review and meta-analysis. JMIR Mhealth Uhealth 2020 Mar 27;8(3):e17776 [FREE Full text] [doi: 10.2196/17776] [Medline: 32217503]
36. Alessa T, Hawley MS, Hock ES, de Witte L. Smartphone apps to support self-management of hypertension: review and content analysis. JMIR Mhealth Uhealth 2019 May 28;7(5):e13645 [FREE Full text] [doi: 10.2196/13645] [Medline: 31140434]
37. Thomopoulos C, Bazoukis G, Ilias I, Tsioufis C, Makris T. Effects of glucose-lowering on outcome incidence in diabetes mellitus and the modulating role of blood pressure and other clinical variables: overview, meta-analysis of randomized trials. J Hypertens 2019 Oct;37(10):1939-1949. [doi: 10.1097/HJH.0000000000002152] [Medline: 31157748]
38. Katsi V, Georgiopoulos G, Vogiatzi G, Oikonomou D, Megapanou M, Skoumas J, et al. Effects of oral and non-insulin injectable antidiabetic treatment in hypertension: a systematic review. Curr Pharm Des 2017;23(25):3743-3750. [doi: 10.2174/1381612823666170519144841] [Medline: 28741456]

39. Stuart EA. Matching methods for causal inference: a review and a look forward. Stat Sci 2010 Feb 01;25(1):1-21 [FREE Full text] [doi: 10.1214/09-STS313] [Medline: 20871802]
40. Thoemmes FJ, Kim ES. A systematic review of propensity score methods in the social sciences. Multivariate Behav Res 2011 Feb 07;46(1):90-118. [doi: 10.1080/00273171.2011.540475] [Medline: 26771582]
41. Ho DE, Imai K, King G, Stuart EA. MatchIt: nonparametric preprocessing for parametric causal inference. J Stat Softw 2011;42(8):1-28. [doi: 10.18637/jss.v042.i08]
42. Kohli N, Peralta Y, Zopluoglu C, Davison M. A note on estimating single-class piecewise mixed-effects models with unknown change points. Int J Behav Dev 2018 Feb 15;42(5):518-524. [doi: 10.1177/0165025418759237]
43. Marre M. Importance of intensive blood pressure control in type 2 diabetes: mechanisms, treatments and current guidelines. Diabetes Obes Metab 2020 Apr 06;22 Suppl 2(S2):33-42. [doi: 10.1111/dom.13975] [Medline: 32250521]
44. Cheung BMY, Li C. Diabetes and hypertension: is there a common metabolic pathway? Curr Atheroscler Rep 2012 Apr;14(2):160-166 [FREE Full text] [doi: 10.1007/s11883-012-0227-2] [Medline: 22281657]
45. Magkos F, Yannakoulia M, Chan JL, Mantzoros CS. Management of the metabolic syndrome and type 2 diabetes through lifestyle modification. Annu Rev Nutr 2009;29:223-256 [FREE Full text] [doi: 10.1146/annurev-nutr-080508-141200] [Medline: 19400751]
46. Nilsson PM, Tuomilehto J, Rydén L. The metabolic syndrome: what is it and how should it be managed? Eur J Prev Cardiol 2019 Dec 26;26(2_suppl):33-46. [doi: 10.1177/2047487319886404] [Medline: 31766917]
47. Chaudhary GMD, Tameez Ud Din A, Chaudhary FMD, Tanveer A, Siddiqui KH, Tameez Ud Din A, et al. Association of obesity indicators with hypertension in type 2 diabetes mellitus patients. Cureus 2019 Jul 01;11(7):e5050 [FREE Full text] [doi: 10.7759/cureus.5050] [Medline: 31511803]
48. Rhee SY, Kim C, Shin DW, Steinhubl SR. Present and future of digital health in diabetes and metabolic disease. Diabetes Metab J 2020 Dec;44(6):819-827 [FREE Full text] [doi: 10.4093/dmj.2020.0088] [Medline: 33389956]
49. Arnett DK, Blumenthal RS, Albert MA, Buroker AB, Goldberger ZD, Hahn EJ, et al. 2019 ACC/AHA guideline on the primary prevention of cardiovascular disease: a report of the American College of Cardiology/American Heart Association task force on clinical practice guidelines. Circulation 2019 Sep 10;140(11):e596-e646 [FREE Full text] [doi: 10.1161/CIR.0000000000000678] [Medline: 30879355]
50. Tsujimoto T, Kajio H. Benefits of intensive blood pressure treatment in patients with type 2 diabetes mellitus receiving standard but not intensive glycemic control. Hypertension 2018 Aug;72(2):323-330. [doi: 10.1161/hypertensionaha.118.11408]
51. Gonzalez JS, Tanenbaum ML, Commissariat PV. Psychosocial factors in medication adherence and diabetes self-management: implications for research and practice. Am Psychol 2016 Oct;71(7):539-551 [FREE Full text] [doi: 10.1037/a0040388] [Medline: 27690483]
52. Shan R, Sarkar S, Martin SS. Digital health technology and mobile devices for the management of diabetes mellitus: state of the art. Diabetologia 2019 Jun 8;62(6):877-887. [doi: 10.1007/s00125-019-4864-7] [Medline: 30963188]
53. American Diabetes Association. Introduction: standards of medical care in diabetes—2021. Diabetes Care 2021 Jan;44(Suppl 1):S1-S2. [doi: 10.2337/dc21-Sint] [Medline: 33298409]
54. Lin P, Grambow S, Intille S, Gallis JA, Lazenka T, Bosworth H, et al. The association between engagement and weight loss through personal coaching and cell phone interventions in young adults: randomized controlled trial. JMIR Mhealth Uhealth 2018 Oct 18;6(10):e10471 [FREE Full text] [doi: 10.2196/10471] [Medline: 30341051]
55. Tighe SA, Ball K, Kensing F, Kayser L, Rawstorn JC, Maddison R. Toward a digital platform for the self-management of noncommunicable disease: systematic review of platform-like interventions. J Med Internet Res 2020 Oct 28;22(10):e16774 [FREE Full text] [doi: 10.2196/16774] [Medline: 33112239]
56. Stark Casagrande S, Ríos Burrows N, Geiss LS, Bainbridge KE, Fradkin JE, Cowie CC. Diabetes knowledge and its relationship with achieving treatment recommendations in a national sample of people with type 2 diabetes. Diabetes Care 2012 Jul;35(7):1556-1565 [FREE Full text] [doi: 10.2337/dc11-1943] [Medline: 22498806]
57. McManus RJ, Little P, Stuart B, Morton K, Raftery J, Kelly J, HOME BP Investigators, et al. Home and online management and evaluation of blood pressure (home bp) using a digital intervention in poorly controlled hypertension: randomised controlled trial. BMJ 2021 Jan 19;372:m4858 [FREE Full text] [doi: 10.1136/bmj.m4858] [Medline: 33468518]
58. Lawson PJ, Flocke SA. Teachable moments for health behavior change: a concept analysis. Patient Educ Couns 2009 Jul;76(1):25-30 [FREE Full text] [doi: 10.1016/j.pec.2008.11.002] [Medline: 19110395]
59. Lustria MLA, Noar SM, Cortese J, Van Stee SK, Glueckauf RL, Lee J. A meta-analysis of web-delivered tailored health behavior change interventions. J Health Commun 2013;18(9):1039-1069. [doi: 10.1080/10810730.2013.768727] [Medline: 23750972]
60. van der Laan LN, Papies E, Hooge I, Smeets P. Goal-directed visual attention drives health goal priming: an eye-tracking experiment. Health Psychol 2017 Jan;36(1):82-90. [doi: 10.1037/hea0000410] [Medline: 27631308]
61. Abbasgholizadeh Rahimi S, Menear M, Robitaille H, Légaré F. Are mobile health applications useful for supporting shared decision making in diagnostic and treatment decisions? Glob Health Action 2017 Jun;10(sup3):1332259 [FREE Full text] [doi: 10.1080/16549716.2017.1332259] [Medline: 28838306]

62. Omboni S, Caserini M, Coronetti C. Telemedicine and m-health in hypertension management: technologies, applications and clinical evidence. High Blood Press Cardiovasc Prev 2016 Sep;23(3):187-196. [doi: 10.1007/s40292-016-0143-6] [Medline: 27072129]

Abbreviations

BG: blood glucose
BP: blood pressure
BMI: body mass index
BPM: blood pressure monitoring
HIPAA: Health Insurance Portability and Accountability Act
LDL: low-density lipoprotein
NBPM: no blood pressure monitoring
T2D: type 2 diabetes

Edited by G Eysenbach; submitted 23.08.21; peer-reviewed by L Kayser; comments to author 14.09.21; revised version received 03.11.21; accepted 31.12.21; published 08.02.22.

Please cite as:
Fundoiano-Hershcovitz , Bacher D, Ritholz MD, Horwitz DL, Manejwala O, Goldstein P
Blood Pressure Monitoring as a Digital Health Tool for Improving Diabetes Clinical Outcomes: Retrospective Real-world Study
J Med Internet Res 2022;24(2):e32923
URL: https://www.jmir.org/2022/2/e32923
doi:10.2196/32923
PMID:35133284

Original Paper

Perceived Acceptability and Experiences of a Digital Psychoeducation and Peer Support Intervention (COPe-support): Interview Study With Carers Supporting Individuals With Psychosis

Rachel Batchelor[1], MSc; Sarah Gulshan[2], MSc; Halpana Shritharan[2], MSc; Elen Williams[3], MSc, MRCGP; Claire Henderson[4], FRCPsych, PhD; Steve Gillard[5], PhD; Luke A Woodham[6], PhD; Victoria Cornelius[7], PhD; Jack Elkes[7], MSc; Jacqueline Sin[5], RMN, PhD

[1]Population Health Research Institute, St George's, University of London, London, United Kingdom
[2]School of Psychology & Clinical Language Sciences, University of Reading, Reading, United Kingdom
[3]Locum GP, London, United Kingdom
[4]Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom
[5]School of Health Sciences, City, University of London, London, United Kingdom
[6]Centre for Technology in Education, St George's, University of London, London, United Kingdom
[7]Imperial Clinical Trials Unit, School of Public Health, Imperial College London, London, United Kingdom

Corresponding Author:
Jacqueline Sin, RMN, PhD
School of Health Sciences
City, University of London
Myddelton Street Building
1 Myddelton Street
London, EC1R 1UW
United Kingdom
Phone: 44 07817027035
Email: jacqueline.sin@city.ac.uk

Abstract

Background: Web-based mental health interventions offer a novel, accessible, and self-paced approach to care delivery to family carers (ie, relatives and close friends who support a loved one with psychosis). We coproduced COPe-support (Carers fOr People with Psychosis e-support), a psychoeducational intervention delivered via an enriched web-based learning environment with network support from professionals and peers. In addition to the rigorous investigation of the effectiveness of COPe-support on the well-being of carers and mental health outcomes, it is imperative to understand the experiences of using the web-based intervention by carers and its associated web-based implementation and facilitation strategies.

Objective: This study aims to explore the experiences of carers and perceived acceptability of COPe-support and its different components, how carers found engagement with COPe-support affected their own well-being and caregiving, and the ideas of carers for improving COPe-support and its delivery to inform any future wider implementation.

Methods: We conducted a qualitative study, individually interviewing 35 carers, following their use of COPe-support for 8 months through a web-based, randomized controlled trial across England. A semistructured guide with open-ended questions was used to explore the experiences of carers and perceived acceptability of the intervention and their ideas to improve the provision. All interviews were conducted remotely through mobile phones or internet communication media, audio recorded and transcribed verbatim. We used a thematic analysis framework to analyze the data.

Results: Three key themes were identified: remote, flexible, and personalized support; impacts on well-being and outlook on caregiving; and future implementation and integration with existing services. Overall, carers found COPe-support a flexible source of knowledge and support from professionals and peers that they could personalize to suit their own needs and convenience. Participants described gaining self-confidence, hope, and a sense of connectivity with others in a similar situation, which helped ameliorate isolation and perceived stigma. Most importantly, COPe-support promoted self-care among the carers themselves. Participants' experiences, use, and activity on COPe-support varied greatly and differed among carers of various ages and levels of computer literacy.

Conclusions: Nearly all participants had a positive experience with COPe-support and supported its wider implementation as a beneficial adjunctive support resource for carers in the future. Any future scale-up of such an intervention needs to consider feedback from carers and suggestions for further improvement. These included having more graphics and audiovisual content materials, improving the navigation, and building in more interactional and customization options to suit various user styles, such as emoji reactions, live web-based chat, opting in and out of updates, and choosing the frequency of reminders. To ensure successful implementation, we should also consider factors pertinent to reaching more carers and integrating the web-based resources with other conventional services.

Trial Registration: International Standard Randomized Controlled Trial Number (ISRCTN) 89563420; https://www.isrctn.com/ISRCTN89563420

International Registered Report Identifier (IRRID): RR2-10.1186/s12888-020-02528-w

(J Med Internet Res 2022;24(2):e27781) doi:10.2196/27781

KEYWORDS

eHealth; family carers; qualitative research; psychosis; peer support; web-based psychoeducation; mobile phone

Introduction

Background

Family members or close friends supporting a loved one affected by psychosis (ie, family or informal carers) play a crucial role in promoting better prognosis and well-being of individuals [1-3]. However, the demands and responsibility of caregiving can make carers vulnerable to physical and mental ill-health [4,5]. Carers need access to psychosocial treatment for knowledge and support to care for their loved ones and to sustain their own well-being [6]. In recent years, with the increasing popularity of digital health interventions targeting a wide range of common mental health symptoms among the general population [7], a few clinical trials investigating such provision for carers of people with psychosis have emerged [8-11]. These web-based interventions tend to be complex multicomponent encompassing psychoeducation (ie, information focused on the health condition and its management) and web-based forums where carers can share emotional support with peers in a *closed* group, for example, COPe-support (Carers fOr People with Psychosis e-support) [12] and Relatives Education And Coping Toolkit [13]. Indeed, psychoeducation on psychosis and related care giving and problem-solving strategies, especially when integrated with peer support among carers, have been identified in systematic reviews as the most desirable ingredients for carer-focused interventions, delivered via the internet or in person [1,14,15]. In previous trials of web-based interventions targeting carers of people with psychosis, psychoeducation was the most common therapeutic approach used. The web-based medium enriched information environment allows carers to self-pick information and advice to suit their own needs and go through them at their own pace [8,11,12]. Psychoeducation and peer support can also target difficulties commonly reported among carers, including isolation, stigma, and uncertainty [16].

Web-based interventions allow for flexible access by carers, minimizing accessibility barriers, such as geographic constraints from needing to be in a particular location and time constraints from juggling multiple roles and responsibilities [17,18]. The web-based medium of delivery also facilitates autonomous use of an individually tailored package of support (ie, carers can choose how and when to use the content at their own convenience) [14,19]. Paradoxically, web-based interventions typically report much lower adherence and completion rates compared with face-to-face interventions, limiting the evidence about their effects [7,14]. Internet support groups and web-based peer forums are often highlighted as desirable features of web-based interventions for promoting social connections and mutual support in mental illness [7,20]. However, their effects, on their own or as part of a complex multimodal intervention, are inconclusive [21-23]. Although users have often identified a peer forum as an engaging element of web-based health interventions [24,25], user characteristics and their use of such forums vary widely [21,26]. Recently, Geramita et al [20] explored the applicability of the 1% rule in a computerized cognitive behavioral therapy platform, which included a patient support group. The 1% rule originated from the web-based marketing literature, suggesting that 1% of participants in web-based communities generate approximately 90% of new content [27]. A computerized cognitive behavioral therapy study [20], among other web-based health intervention trials [24,26], identified that a small number of users (approximately 10%) post most of the content in peer forums, and the remainder mainly observe activity. When considering the use of web-based health tools and services by an individual and subsequent health behavior uptake in general, Powell and Deetjen [26] proposed a new typology. In their study, they identified 6 types of web-based health users (learners, pragmatists, skeptics, worriers, delegators, and adigitals), prompting consideration of the motivation of individuals and orientations behind health-related internet use [26]. Limited evidence to date suggests that high engagement levels with peer forums or discrete elements of complex web-based interventions (eg, information and forums) are associated with better health outcomes or subjective satisfaction or acceptability [7,14,20,28-30]. At the same time, these issues highlight the challenge of implementing complex web-based health interventions that include a peer support forum element with diverse participant profiles and experiences.

Although web-based interventions present a promising opportunity to address a long-standing lack of treatment and support for carers of individuals with psychosis, they can only affect meaningful changes in their users by optimizing their engagement and facilitation strategies to ensure they get the intended benefits. Considering other challenges inherent in developing and evaluating web-based interventions (eg, safety,

personalization, trust, reach, and uptake) [29], it is imperative to embed qualitative process evaluation within web-based intervention trials. While randomized controlled trials (RCTs) are the gold standard study design to establish the clinical effectiveness of an intervention, process evaluation to evaluate the experience of participants and perceived acceptability of the intervention and associated facilitation strategies can identify essential contextual factors in outcomes. For web-based interventions, the contextual factors in question are multiplied, as these interventions are designed to be used autonomously by users in their own homes. Hence, the Medical Research Council complex intervention framework advocates that a thorough process evaluation is needed to understand both the intervention and its implementation process, as experienced by the participants, and to clarify variations in outcomes under contextual influences [31].

Objectives

This qualitative study explores carers' experiences and perceived acceptability of COPe-support and its different components as part of the process evaluation of the COPe-support trial [9,32]. We aim to understand from the carers if and how using COPe-support affected their own well-being and the way they provided care for their loved one. With their experience of using COPe-support, the ideas of carers for improving COPe-support and its delivery were also invited to inform any future wider implementation.

Methods

Research Design and Setting

This study used in-depth individual interviews conducted between February 2019 and October 2020, with participants who had been randomly allocated to use the intervention, after the final follow-up data collection (ie, 8 months after the allocation), as described in the trial protocol [9].

For the RCT of COPe-support, a total of 407 family members or close friends who provided at least weekly support for a loved one affected by psychosis across England were recruited [32]. Over the duration of 2 years (ie, March 2018 to February 2020), 6 cohorts each starting 4 months apart and lasting 8 months were scheduled; when participants consented to participate in the trial, they were allocated to the next cohort scheduled to start [9]. This approach allowed us to group an optimal number of participants (ie, 40-120 participants) established from our earlier systematic reviews into each cohort, which was closed [1,14]. We believe these strategies facilitate peer-group building, thus enhancing the web-based elements of the intervention. Half of the participants were randomly allocated to the intervention arm, that is, access to COPe-support for 8 months, which included being able to post on the peer and expert forums for the initial 4 months (termed the active intervention use period), in addition to usual care. The remaining participants were randomized to receive a web-based noninteractive information bank as an attention-matched control, also with usual care [9].

The Intervention

The web-based intervention COPe-support was coproduced using participatory research methodology as described elsewhere [12]. COPe-support was delivered through a web-based, enriched environment platform that carers could access through a web browser using a computer or a laptop or through an app on smartphones or tablets [9,12]. COPe-support comprises multiple components, including psychoeducation on psychosis and related caring issues, guidance on well-being promotion information and exercises, a *Resource for carers* section signposting to a wide range of external resources weblinks; and 2 web-based forums (one called *Ask the Experts*, where participants could post questions for advice from a panel of experts and the other called *Peer to Peer* for participants to exchange views with one another; see Figures 1-3 for screenshots of COPe-support components). Throughout the study period, a web-based facilitator (an experienced mental health nurse, JS) monitored and moderated all the interactive functions of COPe-support. A weekly email update was sent through the COPe-support platform to all participants for the first 4 months of the study period, which was regarded as the active use period. For security and confidentiality considerations, participants were required to follow a set of ground rules, including using a self-chosen pseudonym and observing confidentiality principles by not sharing any identifying information about themselves and their cared-for person on the COPe-support platform. The web-based intervention platform had an inbuilt use data recording system for log-ins, time spent, and the number of posts made by each participant.

Figure 1. COPe-support (Carers fOr People with Psychosis e-support) home page.

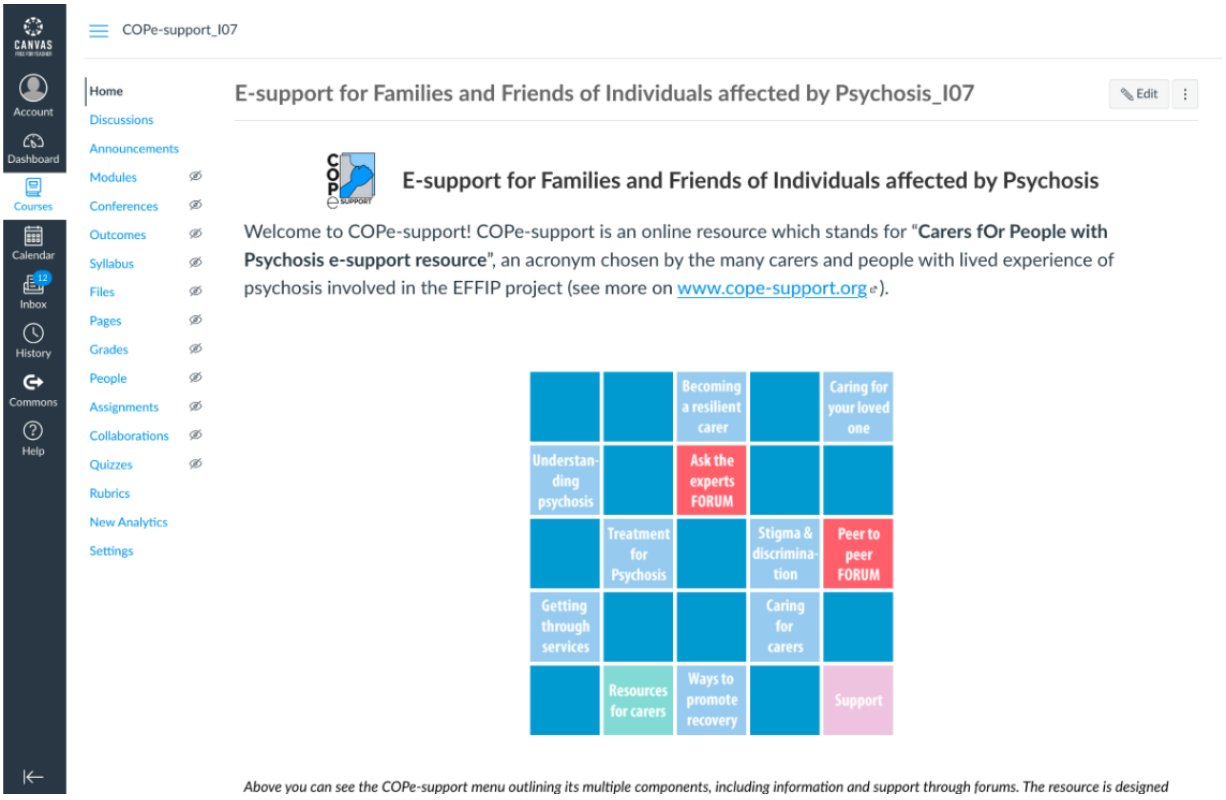


Figure 2. Ask the Experts forum webpage.

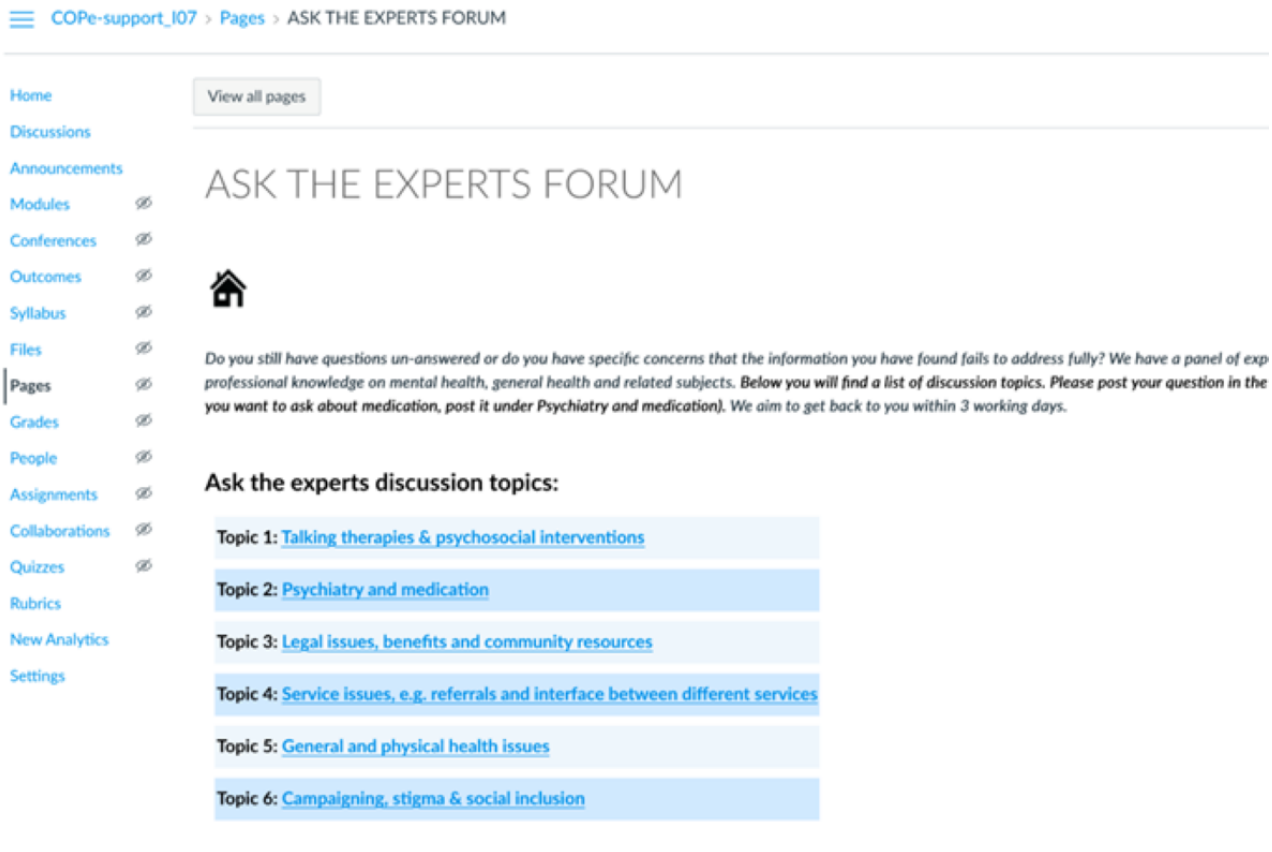
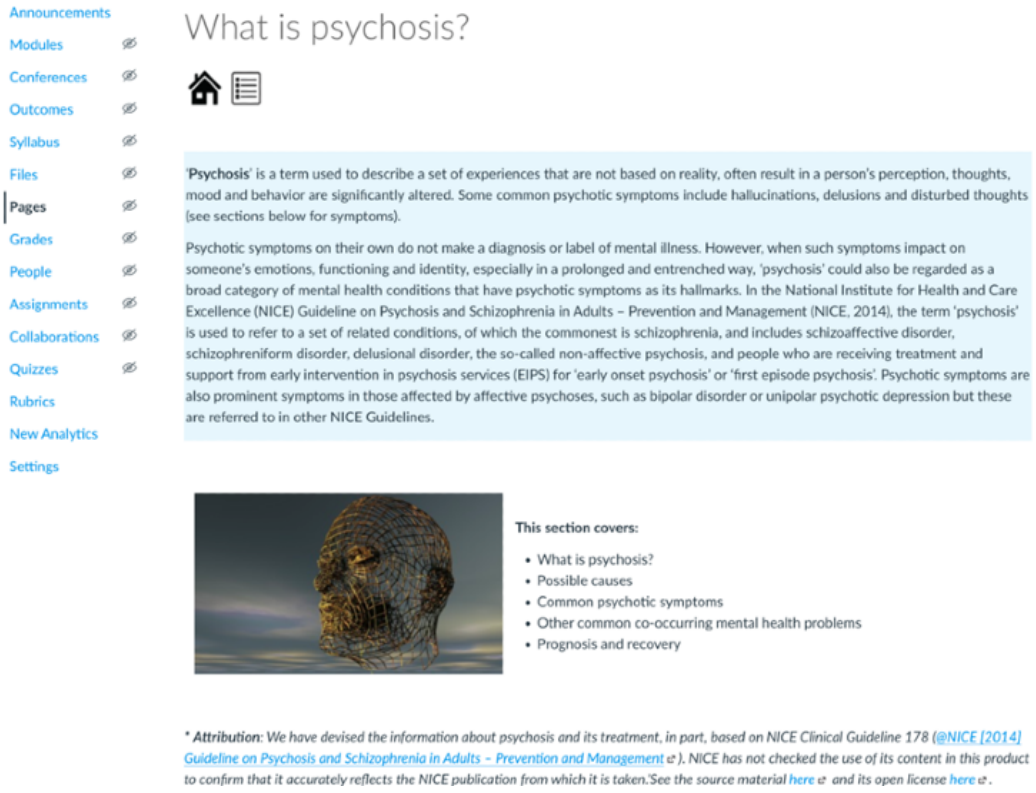


Figure 3. Information on psychosis on COPe-support (Carers fOr People with Psychosis e-support).



Participants

The inclusion criteria for the RCT specified family members, relatives, and close friends who were aged ≥18 years, had at least weekly contact with the cared-for person (in any form ranging from face-to-face to social medical communications), living in England, able to communicate in English in usual web-based communications, and had daily access to the internet, including emails [9]. The inclusion criteria for this qualitative study specified that participants had (1) been randomized to the intervention arm and (2) completed the final RCT follow-up (8 months). Furthermore, purposive sampling was used to identify about 20% of participants in each cohort intervention group from across the *2*-year study period to ensure representation of those with different demographic factors and different levels of use of COPe-support. As previous literature shows that female and White individuals tend to form most of the participants in intervention trials targeting carers for a loved one with psychosis in Western countries [1,3,6,14], we prioritized male carers and those from ethnic minority backgrounds in approaching potential participants. To examine use, we followed the approach by Valentine et al [33] in categorizing participants into levels of use based on the overall number of log-ins by participants to the COPe-support platform over the 4-month active intervention use period. We categorized participants into three use groups as follows: (1) noncompliers, participants who had not activated their log-in or only logged in once; (2) moderate users, those who had logged in between ≥2 and ≤10 times; and (3) high users, those who had logged in >10 times. Within the use categories, we also considered whether participants had forum posts. Thus, participants in the three use groups were further categorized into the following: (1) passive users, participants who did not post, and (2) active users, participants who made at least one forum post, in accordance with previous web-based forum research [34,35]. Participants representing the various demographic considerations and use levels were then contacted via email and invited to participate in an individual interview. A total of 43 participants were invited; 81% (n=35) of the participants agreed, whereas 19% (n=8) of the participants did not respond to the invitation with *2* reminders (their reasons for not responding were not provided).

A total of 35 participants were interviewed and included in the study. The mean age of the participants was 56 years (SD 13 years, range 23-73 years). Most of the participants interviewed were White (29/35, 82%, White British and 2/35, 6%, other White), whereas 6% (2/35) of the participants each described themselves as Asian and Black. Approximately 63% (22/35) of the participants were women, and most (23/35, 66%) cared for a male person. Parents comprised most of the participants (23/35, 66%), followed by partners (7/35, 20%), whereas siblings (2/35, 6%) or close friends (3/35, 8%) formed the remainder. According to the data on their caregiving roles and activities provided by the participants, the mean age of the cared-for persons was 35 years (SD 14 years, range 17-66 years). Just over half (18/35, 51%) of the participants reported that their cared-for persons first became unwell with psychosis <5 years ago, whereas 11% (4/35) of the participants described their loved ones had their first psychotic onset over 20 years previously, and the remainder (13/35, 37%) had been caring for between 5 and just under 20 years. Approximately half (17/35, 49%) of the participants lived with their cared-for person, and 40% (14/35) of carers reported spending over 20 hours per week in caregiving activities. Table 1 provides a summary of the participants' demographic, caregiving, and use data.

Table 1. Summary of participant use and demographic information categorized by use groups.

Users and pseudonym	Cohort (start time)	Sex of carers	Age of carers (years)	Relationship with CfP[a]	Sex of CfP	Age (years) of CfP	Overall weekly log-ins[b]	Overall page-views[c]	Posts made
Noncompliers (who have not logged in or only logged in once throughout the 4 months)									
Passive users (<1 post made)									
Mark	October 18	Male	53	Parent	Male	27	1	3	0
Alexandra_1	February 19	Female	62	Parent	Male	31	1	8	0
Ahmed	February 19	Male	41	Partner	Female	40	1	3	0
Aaron	June 19	Male	23	Partner	Male	22	1	40	0
Sally	October 19	Female	72	Parent	Male	36	__[d]	0	0
Anna	October 19	Female	50	Parent	Female	20	—	0	0
Moderate users (who have logged in between ≥2 and ≤10 times in different weeks)									
Passive users (<1 post made)									
Fern	June 18	Female	57	Sibling	Male	65	4	163	0
Martin	June 18	Male	63	Parent	Female	27	4	83	0
Summer_2	October 18	Female	54	Parent	Male	17	6	238	0
Faye	February 19	Female	68	Parent	Male	38	4	561	0
Alfred	February 19	Male	55	Partner	Female	43	5	263	0
Sam	June 19	Female	71	Parent	Male	42	4	137	0
Polly	February 20	Female	70	Parent	Male	41	8	677	0
Hamish	February 20	Male	55	Parent	Male	30	2	20	0
John	February 20	Male	50	Partner	Female	59	6	482	0
Edward	February 20	Male	33	Partner	Female	30	2	195	0
Active users (≥1 post made)									
Katrina	February 19	Female	62	Sibling	Male	57	2	533	10
Alexandra_2	June 19	Female	72	Parent	Male	32	6	930	15
Alexandra_3	June 19	Female	58	Parent	Male	20	4	602	22
Ben_2	June 19	Male	69	Parent	Male	35	10	974	11
Felix	October 19	Male	42	Stepparent	Male	17	3	226	3
Abbie	October 19	Female	54	Parent	Female	27	5	212	4
Molly	February 20	Female	50	Parent	Male	17	7	248	4
Sophie	February 20	Female	27	Friend	Female	26	3	63	1
Louise	February 20	Female	73	Parent	Male	40	6	206	2
High users (those who have logged in >10 times in different weeks)									
Active users (≥1 post made)									
Matthew	February 19	Male	46	Partner	Female	44	15	1125	1
Flossie	June 18	Female	58	Parent	Male	28	13	654	3
Tony	June 18	Male	43	Partner	Female	41	19	2602	31
Summer_1	October 18	Female	57	Parent	Male	25	11	311	3
Alex	October 18	Female	56	Parent	Male	24	15	554	3
Ben_1	February 19	Male	66	Parent	Male	37	13	715	7
Eleanor	February 19	Female	63	Friend	Female	63	13	967	29
Abby	October 19	Female	67	Partner	Male	66	12	354	4

Users and pseudonym	Cohort (start time)	Sex of carers	Age of carers (years)	Relationship with CfP[a]	Sex of CfP	Age (years) of CfP	Overall weekly log-ins[b]	Overall page-views[c]	Posts made
Maryam	October 19	Female	67	Parent	Female	26	11	1154	10
Imogen	February 20	Female	62	Parent	Female	30	14	1227	7

[a]CfP: cared-for person.

[b]Number of weeks with log-ins across the 4-month active intervention use period.

[c]Total page-views across the 4-month active intervention use period.

[d]Has not activated the log-in.

Multiple participants from each of the different use groups across cohorts were interviewed. All participants were interviewed shortly after their access to the intervention platform ceased (ie, at the 8-month follow-up), although their last access to the platform varied widely. Of the 35 participants, 17% (n=6) participants were classified as noncompliers. All noncompliers were passive users (passive noncompliers). Many participants (19/35, 54%) were classified as moderate users, 53% (10/19) of whom were passive (passive-moderate users) and 47% (9/19) were active (active-moderate users). High users comprised the remaining 28% (10/35) of participants, all of whom were active within the COPe-support forums (active-high users).

Data Collection

All interviews were conducted remotely, suiting the preferences of participants for either phone or internet-facilitated interviews (using Skype [Microsoft] or Microsoft Teams). No face-to-face interviews were used, as all participants had joined the web-based trial of a web-based intervention, with no requirement for in-person contact. Author JS conducted all interviews. Informed written consent was obtained from each participant through the web-based study platform before the interview. At the beginning of each interview, we asked the participants to confirm their consent orally, including for the interview to be audio recorded. All interviews were audio recorded, apart from 3% (1/35) of the participants who opted for their interview recorded by written notes instead.

The interviews followed a topic guide that was devised by the Project Reference Group members, including individuals with lived experiences of psychosis or caring for a loved one with psychosis, who had been involved in developing the intervention [12]. In line with the objectives of this interview study, the interviewer asked open-ended questions to explore the experiences of participants and their views of COPe-support, any specific features of the intervention that they liked or disliked, and the barriers to and facilitators of their access and use of COPe-support, including the facilitation strategies used. The interviewer asked the participants to reflect on their subjective evaluation of the impact of using COPe-support on both themselves and their caregiving experiences. Finally, the interviewer asked the participants for their views and ideas for a plausible wider implementation of COPe-support in the future. The topic guide, which includes the semistructured interview questions and prompts, is presented in Multimedia Appendix 1. Interview times ranged from 14 to 49 minutes and a total of 1117 minutes of data were transcribed.

Data Analysis

The audio recordings were transcribed verbatim. Only transcribed anonymized textual materials were used for the analysis. The data were analyzed in 4 phases using thematic framework analysis [36], with the software NVivo 12 (QSR International) [37]. In accordance with the thematic framework analysis, we commenced the data analysis once the first qualitative interview had been completed and transcribed. To ensure the analysis was grounded in the data and the exploration of the experiences of participants was driven by the emerging results, the interviews and analysis were performed in parallel so that the identified themes and framework of analysis could be tested and validated in latter data.

In the first analysis phase, the authors (JS, S Gulshan, HS, and RB) familiarized themselves with the data by rereading the transcripts and noting interesting aspects. In the second phase, 2 authors (S Gulshan and HS) coded all the data, and a third author (RB) coded 20% of the data independently. The data coded by the third author was selected based on user type and demographics, to ensure that all groups across the full sample were represented. Open (unrestricted) descriptive codes summarizing text segments were applied across the data set. The codes were discussed and reviewed by the authors through several iterations. In the third phase, initial themes and subthemes reflecting broad units of common ideas were formed by grouping relevant codes. These were compared by reviewing the entire data set as well as within individual cases. In the fourth and final phase, the authors (RB, S Gulshan, HS, EW, and JS, all women) cross-referenced, discussed, and clearly defined the themes and subthemes and their interrelated links over several meetings. We used a combined inductive and deductive approach to coding and selecting themes throughout the analysis process [38]. Initially, we used inductive coding, driven by the data (ie, the experience of participants or the way they assigned meaning to their perception of using COPe-support). Nonetheless, as the study aimed to explore the participants' perception of specific elements, functions, and facilitation strategies related to the web-based intervention, we also coded the data deductively with reference to previously reported findings as reported in the literature on wider web-based health interventions and those targeting carers for individuals with a mental illness. These concerned web-based content, forums, facilitation, and perceived safety and security and were explored by questions within our interview topic guide [8,14,24]. Suggested improvements specific to COPe-support were coded deductively using the ideas generated from the views expressed by the participants. Iterative analysis of the transcript

showed that saturation of data was achieved as the final 2 interview transcripts produced no new themes or subthemes [39].

Ethics Approval and Consent to Participate

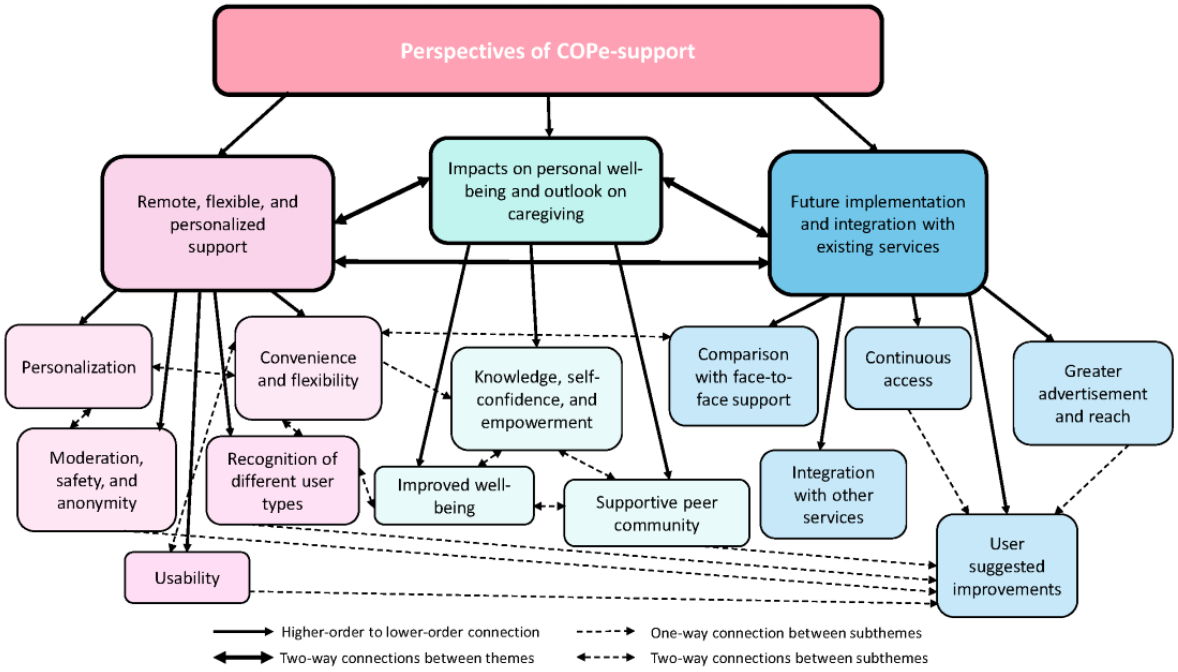
This study, as part of the overall RCT, was reviewed and approved by the South Central—Oxford C Research Ethics Committee (reference: 18/SC/0104) and the Health Research Authority (reference: IRAS 240005). Before study participation, all participants were required to view and give consent on the web to the information provided in the participant information sheet.

Results

Overview

In total, 3 main themes were identified, with each theme divided into subthemes to comprehensively capture the phenomenon explored. The three main themes were as follows: (1) remote, flexible, and personalized support; (2) impacts on well-being and outlook on caregiving; and (3) future implementation and integration with existing services (Figure 4). A brief summary of each theme and subtheme is provided in Multimedia Appendix 2.

Figure 4. Coding tree summarizing the interrelated themes and subthemes. COPe-support: Carers fOr People with Psychosis e-support.



Remote, Flexible, and Personalized Support

This theme covered the experiences and perspectives of carers using COPe-support, with particular regard to navigation, safety, and usability. This theme incorporated 5 subthemes as follows.

Personalization

Carers mostly appreciated that they could choose and focus on specific content on COPe-support, suiting their own circumstances and needs. Several carers also valued the ability to choose their own pseudonyms. Although sharing a common caregiving role, carers recognized that they each have specific interests and needs based on their cared-for person's presentation, treatment, and a range of caregiving factors. For instance, for some carers, information and advice on getting through the benefits system could be a priority at the time, whereas others were after a summary of research evidence of a new treatment:

> *Whereas the stuff from the website was quite helpful and we could tailor it to our own sort of thing. Yes, it was just nice.* [Aaron; male, partner, passive noncomplier]

Some carers discussed a preference for having a greater sense of independence and choice regarding communication in the forums. This included being able to opt in and out of updates and chose the frequency of reminders.

Convenience and Flexibility

Many carers appreciated the convenience of having the information and resources they needed in one place and being able to revisit information and access information at any time and place in day-to-day life. Carers particularly valued the flexibility of their use of COPe-support. This included having autonomy over their use and posting without having to adhere to engagement targets, as well as being able to pick out relevant information at their own pace. Several carers found revisiting and downloading the information for future reference particularly useful:

> *Yes, I mean anything that was easier to download and keep for reference, I mean it's always good to have reference material.* [Ben_2; male, parent, active-moderate user]

Carers particularly appreciated having access to a range of professionals and found it fascinating to receive different perspectives from experts with various experiences and

knowledge. Most carers also appreciated the convenience of expert knowledge on COPe-support. Several considered this novel and felt it addressed the lack of access to experts in existing services for loved ones. Carers valued the opportunity to ask specific questions at any time and received prompt and thought-out answers. A couple of carers noted that this was in contrast to their experiences of feeling rushed within appointments with professionals:

> *I think looking back to seeing the doctors and the psychiatrists you feel a bit rushed and they haven't got time to think about it much but if it's sent as a question you feel someone has taken time to give you an in-depth answer.* [Abby; female, partner, active-high user]

Moderation, Safety, and Anonymity

Being anonymous helped many carers feel more comfortable interacting on forums. Most carers felt that anonymity helped to protect the privacy of their loved one with psychosis and did not affect the community feel on COPe-support:

> *Anonymization probably is quite important because if you are posing questions or comments about your experience as a carer, you inevitably have to talk about that person and they might not like you doing that.* [Maryam; female, parent, active-high user]

Carers particularly appreciated the ground rules (eg, being respectful and not mentioning identifying information) and forum moderation (eg, checking and approving content), providing reassurance that the forums comprised a safe environment. Some carers expressed appreciation and preference for professionals (as used in the trial) over hypothetical carer moderation, providing the professional understood the needs of carers, to help increase the accuracy of information, dispel potentially misguided beliefs, and manage disagreements.

Weekly emails tended to have a positive impact: carers felt they not only served as a reminder for the intervention but also that someone cared. Some carers described themselves as looking forward to or smiling at the emails. Overall, the carers shared a sense that COPe-support was safe and trustworthy, as reflected in the following:

> *It's having a trusted site to look at and knowing that if you put anything on it it's a safe place.* [Abbie; female, parent, active-moderate user]

Usability

Mixed experiences were shared regarding the usability of the COPe-support. Some carers felt confident owing to good computer literacy or previous experience with similar platforms, whereas others described barriers, such as age and poor computer literacy. Carers appreciated accessing COPe-support on different devices (eg, computers, laptops, or mobile phones), with some finding devices with larger screens easier to navigate. Most carers described an adjustment period during which they initially struggled with navigating the interventions but adjusted and grew in confidence over time:

> *I did start to get a bit more used to [navigating] after a while but to begin with I did find it complicated.* [Summer_1; female, parent, active-high user]

Recognition of Different User Types

Many carers expressed awareness of different user types on COPe-support. Carers tended to distinguish between enthusiastic (active) users whose names frequently appeared within forums and other (passive) users who tended to observe. Some active users reported focusing on the peer and expert forums and felt these aspects in themselves made COPe-support *powerful* (Eleanor; female, friend, active-high user and Alexandra_3; female, parent, active-moderate user). A couple of active users were unaware of being particularly active and felt unsure of how many people read their posts. Most passive users were aware that they had not posted and found reading what others had to say useful and knowing the forums were there if needed comforting in itself:

> *I guess there are some people who are going to be very active on there and discuss things a lot and then there are going to be people who are very quiet on there...It doesn't mean to say that they're not taking it all in and getting something from it...I also regret slightly now that I wasn't a bit more active at the same time.* [Summer_2; female, parent, passive-moderate user]

Impacts on Well-being and Outlook on Caregiving

Three subthemes represented the impacts on well-being and outlook on caregiving experienced from participating in COPe-support.

Knowledge, Self-confidence, and Empowerment

Many caregivers felt that COPe-support provided comprehensive, relevant, and helpful information across various important topics. New carers found the information especially suitable for their first time learning about psychosis. For others with existing knowledge, the information supplemented the resources they had previously accessed. Although some felt COPe-support had enhanced their knowledge and skills enough to not require further support, others appreciated the signposting to other sources and local and national services to further support their loved ones:

> *I've got 99% certain I will either get signposted in the right direction or find what I want rather than Googling and going through different websites and trying to find the same information.* [Faye; female, parent, passive-moderate user]

Some carers felt that the information was quite generic, outdated, and repetitive. Moreover, some newer carers initially found the amount of information overwhelming, although they reported adjusting and learning over time:

> *It's also a strength is the fact that once you are in the program you realize just how comprehensive and detailed it actually is and that this could be a bit daunting initially for people signing up.* [Eleanor; female, friend, active-high user]

Most carers felt that the tone of the experts was just right: not pressurizing or patronizing, yet empathetic, respectful, and comforting. Understandable language (eg, no acronyms, abbreviations, and technical terms) was also used to explain complex information in an understandable way. However, several carers felt that the answers were sometimes generic or vague, although they appreciated that the experts were not aware of the full situation of their loved ones and still found the suggestions helpful. The information provided and knowledge gained subsequently empowered them to seek further conversations with mental health professionals caring for loved ones, as expressed by a participant:

> *They've not been able to provide really specific answers sometimes because obviously they don't know our situation but the fact of it is they've been able to signpost or suggest something that you maybe hadn't thought of.* [Alexandra_3; female, parent, active-moderate user]

Some carers discussed the lack of preparation for caring roles and the ongoing self-doubt surrounding doing the right thing or supporting their loved one in a helpful way. These carers felt by gaining information, resources, and knowledge on COPe-support had better equipped them and also improved their self-confidence in their ability as carers:

> *I'm sure it's given me more confidence as a carer because I've got more information and that also becomes a part of how I care for my daughter and talk to the family and others as well.* [Maryam; female, parent, active-high user]

Supportive Peer Community

One significant benefit identified was a sense of belonging to a supportive peer community, without ever seeing or knowing one another. Many carers discussed feelings of loneliness and isolation experienced by them. Reading the resources and forums showed carers that others were experiencing similar and relatable difficulties, helping them feel less alienated, isolated, and detached:

> *Sometimes when you are a carer you think you are alone. When you go to these groups or you do these things, you realise you are not. It makes a difference.* [Anna; female, parent, passive noncomplier]

Carers also reported feeling more connected and having a sense of solidarity and unity with others to proceed on the caring journey. Carers valued having a sense of community, group alliance, and connection, which naturally arose from sharing similar experiences and challenges and feeling mutually understood, something they often lacked in their own lives. This was made explicit by the following:

> *Being able to see that people are getting some support and that it normalizes the issues that we don't talk about.* [Abbie; female, parent, active-moderate user]

Most carers valued being linked with other carers, especially new carers who felt shell-shocked and craved speaking to others in a nonjudgmental environment. Many carers appreciated the opportunity to learn from peers, including practical tips, advice, and awareness of differing carer experiences. Some also valued the opportunity to help other carers and the positive feelings that came with that:

> *But in the main I found the whole thing quite helpful especially for the first month or so when I could see or read about everyone else's problems and some were similar to mine and some of the advice they gave if you know what I mean.* [Ben_1; male, parent, active-high user]

In addition to creating a community and reducing loneliness, many carers noted that reading posts from other carers also helped normalize and validate their feelings and experiences. The intervention content and forums also helped to normalize concerns, fears, and often stigmatized psychosis-related topics that carers often found difficult to talk to people in their personal lives about. Such normalization and validation subsequently helped carers feel less overwhelmed:

> *Yes, I think I found it really helpful as well because some of the ways that it was designed with the different subjects helped as well to make me think oh yes well this experience I'm having is normal, which is like there was, how it was set up the program it had stigma.* [Eleanor; female, friend, active-high user]

Some carers felt that the expert and peer support forums provided hope, particularly in instances where carers were able to provide lived accounts and reassurance of particular aspects and situations improving over time. Some carers especially valued reminders that their loved one is still their loved one and reflected that kind words provided *light*, led to a feeling of hope:

> *In some respects, it made me feel a bit better because other people are going through not completely the same as me but very similar as me and they've managed to get through it, etc.* [John; male, partner, passive-moderate user]

Improved Well-being

Carers recognized that the COPe-support was specifically designed for them. Some carers discussed how COPe-support not only provided support for the well-being of their loved ones but also their own. This included recognizing the importance of supporting their own needs, focusing on self-care, and fostering healthier routines, such as improving their diet, fitness, and sleep hygiene:

> *It was just really, really helpful to learn how I can manage my well-being in terms of trying to support myself in terms of trying to help the person I'm caring for...like I said it has made a really big difference to my well-being and my partner's well-being and it has been a lifeline.* [Edward; male, partner, passive moderate user]

> *I have actually changed my eating this last few months as a direct result of the site, so that's quite something.* [Alexandra_3; female, parent, active-moderate user]

Even the concept that COPe-support had been designed specifically for carers helped carers recognize their support needs were valid and acknowledged, reducing guilt associated

with personal help-seeking. Some carers described how COPe-support had provided personal space and time to reflect on their personal journey as carers, get more in touch with their emotions, and listen to the reflections of others:

> *Even just using the questionnaires at times were good for me because it made me sit and focus a little bit on where things were at...and actually think about how I was feeling.* [Alexandra_1; female, parent, passive noncomplier]

Future Implementation and Integration With Existing Services

The following subthemes reflect the perspectives of carers surrounding the future implementation of COPe-support and integration with existing services. This includes suggested improvements for COPe-support.

Comparison With Face-to-face Support

Compared with face-to-face support for carers, the perspectives of COPe-support were mixed. Although some expressed a preference for traditional means of delivery, others preferred web-based platforms and ideally a blended approach. Barriers to face-to-face support, including geographic factors, family life, funding and time constraints, and the benefits of web-based delivery in minimizing these barriers were discussed by some carers. Other carers considered barriers to web-based interventions, including age and a desire to personally meet carers and be able to sit with others going through similar situations. This is expressed as follows:

> *And I think that e-support is definitely a very, very useful, well it's a very good use of technology for people who have computers or phones and have the confidence to access stuff. You can't beat that one-to-one when you need it, you can't beat that.* [Faye; female, parent, passive-moderate user]

Integration With Other Services

Several carers commented that, given the funding restrictions on existing services, implementing COPe-support could *only be a benefit*. Some carers highlighted that participating in COPe-support addressed their concerns surrounding interventions for carers and motivated use of other services, such as face-to-face groups and courses for carers. Although some felt the support they had received through COPe-support was sufficient for their needs, others emphasized that COPe-support should serve as an adjunct to existing services rather than a replacement:

> *It also encouraged me to join a carers and coping course...I think it's made me question why I would find it so hard...to sit in a group with other people and hear about what's been happening to them, so yes I'm definitely looking forward to going to a six-week course at the end of this month.* [Summer_2; female, parent, passive-moderate user]

Continuous Access

Perspectives on the length of time to access COPe-support were mixed. Some felt they had received access for just the right amount of time to remain engaged and gain optimal benefits as a carer. However, some desired a longer use time. Several carers highlighted that as caring can be a long and complex journey, it would be reassuring to be able to revisit information and know they would be able to use it and have instant access to support in the future if new challenges arise (ie, dip in and out):

> *People have different periods of crisis. You would not want to have the sense of support suddenly be taken away.* [Martin; male, parent, passive-moderate user]

To allow for continuous access to carers' needs, some suggested being able to self-refer back into the intervention if necessary or have continual access and be able to opt out when they felt they had used it enough:

> *It was very good, too good; hence I asked if I can enroll again...it was a lifeline for me...COPe-support came along and gave me all the help and support I've ever wanted.* [Summer_1; female, parent, active-moderate user]

Greater Advertisement and Reach

Some carers reflected on coming across COPe-support *by chance* and emphasized a need for greater advertisement to reach more carers if it was to be rolled out widely in the future. Several advertising and promotion routes have been suggested, including local authorities and social services, charities, general physician surgeries, existing services for carers and trust websites, noticeboards, and newsletters. Awareness among health and social care professionals was also noted as important, with potential screening for the well-being of carers and onward signposting to COPe-support recommended. Suggestions for ways to reach carers include the following:

> *When you roll it out into various Trusts and it goes further that's where it needs to be as well. There are a number of options there.* [Mark; male, parent, passive noncomplier]

> *They always ask at the GP surgery when you register or every so often they'll say are you caring for anyone and it could be quite helpful to maybe signpost it at that point.* [Aaron; male, partner, passive noncomplier]

User Suggested Improvements

Many carers have proposed improvements for COPe-support. Some were about the way information was presented, which some found, at times, overwhelming and off-putting (ie, *too much on the screen sometimes*). To reduce confusion, fewer chunks of text and more graphics and visual aids or *see more* dropdown options were recommended:

> *I suppose what I'm trying to say is even a little, you need to have something...if you want to do mindfulness it needs to have a little picture, it needs to be more visually stimulating.* [Molly; female, parent, active-moderate user]

With regard to forum communications, although some described freely writing open posts as cathartic, several carers reflected on an emotional burden arising from posts. At times, carers found posts distressing to read and that they could generate

worries. Thus, providing a general warning of content causing potential distress and content warnings for particular comments was recommended:

> *There were things where...you know, there were things that triggered me to think about things and thought this maybe something worth sharing.* [Matthew; male, partner, active-high user]

Moreover, some caregivers desired more ongoing conversations. Thus, suggestions for a chatroom or befriender element were made by a few carers to help build stronger connections. Some carers reported it was hard to relate to others given different life circumstances (eg, having several children to care for too). Hence, a couple of carers recommended brief profiles with basic, yet nonidentifying, information (eg, sex, caring responsibilities, relationship, or living situation) to provide advice and support, as well as seeking relatable content. However, when certain forum topics received a good number of posts, one common problem that arose was having to go forward and backward among pages and scrolling excessively to see forum comments. This was described by an active user as follows:

> *I remember the format of the message threads when you had five or six interactions or replies on the same thread it becomes almost impossible to read on the phone because you have to scroll down and the indentation starts going to the right.* [Tony; male, partner, active-high user]

Hence, for navigating the forums and the COPe-support content overall, *frequently viewed* and *recently viewed* buttons were recommended by some carers. Some would also like to be able to choose which posts on the forum to expand. Most found the instructions for navigating COPe-support clear, although some would have appreciated an opt-in for 1:1 guidance.

Finally, to encourage engagement, some carers noted that they would have appreciated some additional prompting after periods of inactivity. Several passive users regretted not using the forums more and reported barriers to posting, including their busyness, mental state, difficulties expressing their feelings, worries surrounding sharing with unknown people, and experiencing hesitation and self-doubt. Some carers suggested having rolling discussion topics and implementing alternative options (eg, emoji reactions) to facilitate forum engagement:

> *If that [thumbs up or other emojis for acknowledgement] feature had been available and I'd seen a couple of thumbs up to the things I'd posted I think that would have been great...And maybe that's a stepping stone as well they start by just a few reactions, emoji reactions and then it's small steps. They can do that the first time and then maybe the next time they will write a few words.* [Felix; male, stepparent, active-moderate user]

Discussion

Principal Findings

This study aimed to explore the following: (1) carers' experiences and perceived acceptability of COpe-support and its different components; (2) how they found engagement with COPe-support affected their own well-being and caregiving; and (3) ideas of carers for improving COPe-support and its delivery to inform any future wider implementation. Notably, this qualitative study is one of the first to explore the experiences of carers of individuals with psychosis by using an entirely web-based psychoeducation and peer support intervention, coproduced by carers and people with experiential expertise. Experiences of participants were predominantly positive with COPe-support, and carers identified a range of benefits from using the intervention. Nonetheless, the carers highlighted some key areas of improvement. Overall, three themes were identified, each addressing one of the objectives of the study as follows: (1) remote, flexible, and personalized support; (2) impacts on well-being and outlook on caregiving; and (3) future implementation and integration with existing services.

Overall, the subjective experiences of COPe-support among carers were positive. In addition to the web-based gains provided by COPe-support, such as improved accessibility, flexibility, and anonymity, participants also reported that the intervention was beneficial in providing access to a rich repertoire of credible information [8,11,12] and fostering personal development by enhancing their self-confidence and understanding. Our results indicate that COPe-support was perceived as a crucial resource to reinforce feelings of empowerment in carers while reducing their sense of isolation. COPe-support also prompted carers to prioritize their own well-being. These impacts motivated some carers to access further support and engage more with professionals, indicating additional long-term benefits [1,3,6,14,15].

Notably, our themes and subthemes should be recognized as a set of interconnected and interacting constructs to be considered in the overall design (eg, content) and facilitation (eg, moderation) of web-based interventions, such as COPe-support [8]. For instance, carers would only enjoy interacting on the web-based forums, provided they felt safe and supported through specific implementation strategies. Carers would be less likely to see the essential intervention contents should access and navigation be less than facilitative.

Moreover, similar to earlier studies on web-based interventions with a forum component [20-23,33], we found that the use by carers in terms of numbers of posts and log-ins does not always align with their perceived acceptability and usefulness of the intervention. Although the carers who actively initiated posts themselves were eager to see more exchanges on the forums, many others found benefits in being passive observers. Some carers identified that anonymous participation on web-based forums allowed them not to feel pressurized to participate in a certain fashion as in a face-to-face group setting. Many carers described finding resonance, connections, and solidarity from the peer and expert forums without making a post themselves, although some identified that they would have made posts if given more time or if a specific question came up.

Future Directions

The experience of participants seemed partly dependent on factors, such as their own demographic profile (eg, length of time as a carer and age) and preferences for particular delivery

formats and computer literacy as highlighted in previous studies [14,26,40,41]. It is imperative to incorporate these perspectives in considering how best to further refine COPe-support and its facilitation. Upon future implementation, several advertisement routes were recommended to increase the reach of COPe-support, as well as a need for greater awareness among professionals who have contact with carers. In line with previous research [41], the need for more proactive approaches from professionals and services to identify and refer carers were highlighted, such as potential screening for the well-being of carers and signposting to COPe-support [32].

In any future rollout of COPe-support, it is imperative to consider the revision and refinement of the content as much as the facilitation of the minimally guided web-based intervention holistically to keep the participants engaged, to induce the anticipated impact [7]. Further scaling-up implementation of COPe-support and similar interventions also needs to carefully consider what constitutes the optimal group size and setup for a multicomponent web-based object, including closed forums catering for numerous users with varying use and participation profiles and a more flexible time frame to suit the ongoing needs of carers. Some participants also highlighted a desire for blended services—that is, COPe-support being adjunctive to, rather than a replacement of, in-person support. Indeed, a blended approach could foster the discussed benefits of both web-based and face-to-face support, as well as provide carers with options to cater for their needs and preferences.

Strengths and Limitations

We considered the sample of 35 carers interviewed for this study as a strength, as this contributed to a wide variation in user experience and use from carers with different relationships with individuals with psychosis and in different caregiving situations. Having multiple researchers to independently code and analyze the rich data led to unanimous results and increased the rigor and reflexivity of the study [36,42]. The study results allow us to understand how carers engaged with COPe-support, what helped or hindered their engagement, and how using it affected themselves. Through these results, we underscored that carers' experiences of COPe-support were shaped by a range of demographic and web-based health literacy factors, in addition to the intervention design and delivery itself. To ensure that users obtain the intended benefits of the COPe-support in the future rollout, it is imperative to consider how best to engage a wide variety of users to use all its essential ingredients [7,14].

This study had several limitations. Although we aimed to interview carers after completion of outcome data collection at 8 months, some carers had stopped using COPe-support earlier than the study duration and hence found it difficult to recall their experience with the intervention in detail. Although we strived to invite participants with low use and those from ethnic minority backgrounds for the interviews, such populations remained underrepresented (in the overall trial and this study) [43,44]. Our interviewees may have been positively biased in their views surrounding the intervention and study. It could be valuable to extend future work to explore reasons for nonenrollment among potential participants within services where the intervention was advertised, yet they chose not to take part.

Conclusions

Overall, this qualitative interview study captured the experiences of carers of using the web-based intervention COPe-support. The variation in responses among active and passive users captured the carer's perception of COPe-support. Notably, support and engagement with peers and experts were appreciated for meeting and validating the needs of carers, and the importance of usability ease, personalization, convenience, and safety were discussed. Further work is required to develop COPe-support based on these suggestions and explore the steps for optimal implementation.

Acknowledgments

The authors thank all the carers who participated in the study. The authors thank all the input and contributions from the members of the Project Reference Group, which provided an overview of the research program. The members include Ellen Harris; Jacqueline Marks; Angela Ryan; Storm Ryan; Lana Samuels; Clive Travis, PhD; Leigh Wallbank; and EW.

We thank all the members who helped to run the Ask the Experts forum. They include Ana Maria Corredor Collazos; David Coughlin; Dr Ranjita Dhital; Sian Evans; Dr Ban Haider; Julia Heathcote; Dr Claire Henderson; Dr Sarah Mansfield; Dr Aileen O'Brien; Mona Qassim; Juliet Sserunkuma; Clive Travis, PhD; and EW.

To recruit participants into the study, the project team acknowledges the support of the National Institute for Health Research through the Clinical Research Network (Division 4).

This trial, as part of a larger research program entitled E-support for Families and Friends of Individuals Affected by Psychosis, is funded by the National Institute for Health Research under its Postdoctoral Research Fellowship (awarded to JS, reference: PDF-2015-08-035). The study was peer reviewed by the funding bodies.

The views expressed in this publication are those of the authors and not necessarily those of the National Health Service, the National Institute for Health Research, or the Department of Health and Social Care. The funder National Institute for Health Research had no involvement in the study design; collection, analysis, or interpretation of data; writing of the manuscript; or decision to submit the manuscript for publication.

Authors' Contributions

JS conceived the study and initiated the study design with supervision from CH and S Gillard. RB, JS, CH, S Gillard, LAW, S Gulshan, and HS further developed the study design and its implementation. RB, S Gulshan, HS, EW, and JS analyzed and interpreted the results. RB and JS drafted the paper, and S Gulshan, HS, and EW supported its further revision. All authors read and approved the final manuscript. JS is the grant holder.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview topic guide.
[DOCX File , 24 KB - jmir_v24i2e27781_app1.docx]

Multimedia Appendix 2
Summary of themes and subthemes.
[DOCX File , 20 KB - jmir_v24i2e27781_app2.docx]

References

1. Sin J, Gillard S, Spain D, Cornelius V, Chen T, Henderson C. Effectiveness of psychoeducational interventions for family carers of people with psychosis: a systematic review and meta-analysis. Clin Psychol Rev 2017 Aug;56:13-24 [FREE Full text] [doi: 10.1016/j.cpr.2017.05.002] [Medline: 28578249]
2. The Abandoned Illness: a report by the Schizophrenia Commission. Mental Health Partnerships. URL: https://mentalhealthpartnerships.com/resource/the-abandoned-illness-a-report-by-the-schizophrenia-commission/ [accessed 2022-01-18]
3. Pharoah F, Mari J, Rathbone J, Wong W. Family intervention for schizophrenia. Cochrane Database Syst Rev 2010 Dec 08(12):CD000088 [FREE Full text] [doi: 10.1002/14651858.CD000088.pub2] [Medline: 21154340]
4. Singleton N, Maung N, Cowie A, Sparks J, Bumpstead R, Meltzer H. Mental Health of Carers. London: Stationery Office; 2002.
5. Smith L, Onwumere J, Craig T, McManus S, Bebbington P, Kuipers E. Mental and physical illness in caregivers: results from an English national survey sample. Br J Psychiatry 2014 Sep;205(3):197-203. [doi: 10.1192/bjp.bp.112.125369] [Medline: 25061119]
6. Yesufu-Udechuku A, Harrison B, Mayo-Wilson E, Young N, Woodhams P, Shiers D, et al. Interventions to improve the experience of caring for people with severe mental illness: systematic review and meta-analysis. Br J Psychiatry 2015 Apr;206(4):268-274. [doi: 10.1192/bjp.bp.114.147561] [Medline: 25833867]
7. Sin J, Galeazzi G, McGregor E, Collom J, Taylor A, Barrett B, et al. Digital interventions for screening and treating common mental disorders or symptoms of common mental illness in adults: systematic review and meta-analysis. J Med Internet Res 2020 Sep 02;22(9):e20581 [FREE Full text] [doi: 10.2196/20581] [Medline: 32876577]
8. Lobban F, Akers N, Appelbe D, Chapman L, Collinge L, Dodd S, et al. Clinical effectiveness of a web-based peer-supported self-management intervention for relatives of people with psychosis or bipolar (REACT): online, observer-blind, randomised controlled superiority trial. BMC Psychiatry 2020 Apr 14;20(1):160 [FREE Full text] [doi: 10.1186/s12888-020-02545-9] [Medline: 32290827]
9. Sin J, Henderson C, Cornelius V, Chen T, Elkes J, Woodham LA, et al. COPe-support - a multi-component digital intervention for family carers for people affected by psychosis: study protocol for a randomized controlled trial. BMC Psychiatry 2020 Mar 17;20(1):129 [FREE Full text] [doi: 10.1186/s12888-020-02528-w] [Medline: 32183772]
10. Glynn SM, Randolph ET, Garrick T, Lui A. A proof of concept trial of an online psychoeducational program for relatives of both veterans and civilians living with schizophrenia. Psychiatr Rehabil J 2010;33(4):278-287. [doi: 10.2975/33.4.2010.278.287] [Medline: 20374986]
11. Rotondi AJ, Anderson CM, Haas GL, Eack SM, Spring MB, Ganguli R, et al. Web-based psychoeducational intervention for persons with schizophrenia and their supporters: one-year outcomes. Psychiatr Serv 2010 Nov;61(11):1099-1105. [doi: 10.1176/ps.2010.61.11.1099] [Medline: 21041348]
12. Sin J, Henderson C, Woodham LA, Sesé Hernández A, Gillard S. A multicomponent eHealth intervention for family carers for people affected by psychosis: a coproduced design and build study. J Med Internet Res 2019 Aug 06;21(8):e14374 [FREE Full text] [doi: 10.2196/14374] [Medline: 31389333]
13. Honary M, Fisher NR, McNaney R, Lobban F. A web-based intervention for relatives of people experiencing psychosis or bipolar disorder: design study using a user-centered approach. JMIR Ment Health 2018 Dec 07;5(4):e11473 [FREE Full text] [doi: 10.2196/11473] [Medline: 30530457]

14. Sin J, Henderson C, Spain D, Cornelius V, Chen T, Gillard S. eHealth interventions for family carers of people with long term illness: a promising approach? Clin Psychol Rev 2018 Mar;60:109-125 [FREE Full text] [doi: 10.1016/j.cpr.2018.01.008] [Medline: 29429856]
15. Sin J, Norman I. Psychoeducational interventions for family members of people with schizophrenia: a mixed-method systematic review. J Clin Psychiatry 2013 Dec;74(12):e1145-e1162. [doi: 10.4088/JCP.12r08308] [Medline: 24434103]
16. Onwumere J, Shiers D, Chew-Graham C. Understanding the needs of carers of people with psychosis in primary care. Br J Gen Pract 2016 Aug;66(649):400-401 [FREE Full text] [doi: 10.3399/bjgp16X686209] [Medline: 27481960]
17. Hazell CM, Jones CJ, Pandey A, Smith HE. Barriers to recruiting and retaining psychosis carers: a case study on the lessons learned from the Caring for Caregivers (C4C) trial. BMC Res Notes 2019 Dec 17;12(1):810 [FREE Full text] [doi: 10.1186/s13104-019-4832-9] [Medline: 31847874]
18. McCann T, Lubman D, Clark E. First-time primary caregivers' experience accessing first-episode psychosis services. Early Interv Psychiatry 2011 May;5(2):156-162. [doi: 10.1111/j.1751-7893.2010.00246.x] [Medline: 21352509]
19. Sherifali D, Ali MU, Ploeg J, Markle-Reid M, Valaitis R, Bartholomew A, et al. Impact of internet-based interventions on caregiver mental health: systematic review and meta-analysis. J Med Internet Res 2018 Jul 03;20(7):e10668 [FREE Full text] [doi: 10.2196/10668] [Medline: 29970358]
20. Geramita EM, Herbeck Belnap B, Abebe KZ, Rothenberger SD, Rotondi AJ, Rollman BL. The association between increased levels of patient engagement with an internet support group and improved mental health outcomes at 6-month follow-up: post-hoc analyses from a randomized controlled trial. J Med Internet Res 2018 Jul 17;20(7):e10402 [FREE Full text] [doi: 10.2196/10402] [Medline: 30021711]
21. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. BMJ 2004 May 15;328(7449):1166 [FREE Full text] [doi: 10.1136/bmj.328.7449.1166] [Medline: 15142921]
22. Ali K, Farrer L, Gulliver A, Griffiths KM. Online peer-to-peer support for young people with mental health problems: a systematic review. JMIR Ment Health 2015;2(2):e19 [FREE Full text] [doi: 10.2196/mental.4418] [Medline: 26543923]
23. Griffiths KM, Calear AL, Banfield M, Tam A. Systematic review on Internet Support Groups (ISGs) and depression (2): what is known about depression ISGs? J Med Internet Res 2009 Sep 30;11(3):e41 [FREE Full text] [doi: 10.2196/jmir.1303] [Medline: 19793718]
24. Sin J, Henderson C, Norman I. Usability of online psychoeducation for siblings of people with psychosis. Int J Technol Assess Health Care 2014 Oct;30(4):374-380. [doi: 10.1017/S0266462314000488] [Medline: 25394550]
25. Alvarez-Jimenez M, Alcazar-Corcoles M, González-Blanch C, Bendall S, McGorry P, Gleeson J. Online, social media and mobile technologies for psychosis treatment: a systematic review on novel user-led interventions. Schizophr Res 2014 Jun;156(1):96-106. [doi: 10.1016/j.schres.2014.03.021] [Medline: 24746468]
26. Powell J, Deetjen U. Characterizing the digital health citizen: mixed-methods study deriving a new typology. J Med Internet Res 2019 Mar 05;21(3):e11279 [FREE Full text] [doi: 10.2196/11279] [Medline: 30835238]
27. The 90-9-1 rule for participation inequality in social media and online communities. Nielsen Norman Group. URL: https://www.nngroup.com/articles/participation-inequality/ [accessed 2022-01-18]
28. Murray E, Burns J, May C, Finch T, O'Donnell C, Wallace P, et al. Why is it difficult to implement e-health initiatives? A qualitative study. Implement Sci 2011 Jan 19;6(1):6 [FREE Full text] [doi: 10.1186/1748-5908-6-6] [Medline: 21244714]
29. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating digital health interventions: key questions and approaches. Am J Prev Med 2016 Nov;51(5):843-851 [FREE Full text] [doi: 10.1016/j.amepre.2016.06.008] [Medline: 27745684]
30. Murray E, Treweek S, Pope C, MacFarlane A, Ballini L, Dowrick C, et al. Normalisation process theory: a framework for developing, evaluating and implementing complex interventions. BMC Med 2010;8:63 [FREE Full text] [doi: 10.1186/1741-7015-8-63] [Medline: 20961442]
31. Developing and evaluating complex interventions. Medical Research Council. URL: https://mrc.ukri.org/documents/pdf/complex-interventions-guidance/ [accessed 2022-01-18]
32. Sin J, Elkes J, Batchelor R, Henderson C, Gillard S, Woodham LA, et al. Mental health and caregiving experiences of family carers supporting people with psychosis. Epidemiol Psychiatr Sci 2021 Jan 08;30:e3 [FREE Full text] [doi: 10.1017/S2045796020001067] [Medline: 33416043]
33. Valentine L, McEnery C, O'Sullivan S, Gleeson J, Bendall S, Alvarez-Jimenez M. Young people's experience of a long-term social media-based intervention for first-episode psychosis: qualitative analysis. J Med Internet Res 2020 Jun 26;22(6):e17570 [FREE Full text] [doi: 10.2196/17570] [Medline: 32384056]
34. Abel F, Bittencourt II, Costa E, Henze N, Krause D, Vassileva J. Recommendations in online discussion forums for e-learning systems. IEEE Trans Learn Technol 2010 Apr;3(2):165-176. [doi: 10.1109/tlt.2009.40]
35. Mustafaraj E, Bu J. The visible and invisible in a MOOC discussion Forum. In: Proceedings of the Second (2015) ACM Conference on Learning @ Scale. 2015 Presented at: L@S '15: Proceedings of the Second (2015) ACM Conference on Learning @ Scale; Mar 14-18, 2015; Vancouver, Canada URL: https://dl.acm.org/doi/10.1145/2724660.2728691 [doi: 10.1145/2724660]

36. Spencer L, Ritchie J, Ormston R, O'Connor W, Barnard M. Analysis: principles and processes. In: Qualitative Research Practice. London: SAGE Publications Ltd; 2014.
37. NVivo qualitative data analysis software 2018. NVivo. URL: https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home [accessed 2022-01-18]
38. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. Int J Qual Methods 2016 Nov 29;5(1):80-92. [doi: 10.1177/160940690600500107]
39. Guest G, Bunce A, Johnson L. How many interviews are enough? Field Methods 2016 Jul 21;18(1):59-82 [FREE Full text] [doi: 10.1177/1525822x05279903]
40. Köhle N, Drossaert CH, Oosterik S, Schreurs KM, Hagedoorn M, van Uden-Kraan CF, et al. Needs and preferences of partners of cancer patients regarding a web-based psychological intervention: a qualitative study. JMIR Cancer 2015 Dec 29;1(2):e13 [FREE Full text] [doi: 10.2196/cancer.4631] [Medline: 28410157]
41. Carduff E, Finucane A, Kendall M, Jarvis A, Harrison N, Greenacre J, et al. Understanding the barriers to identifying carers of people with advanced illness in primary care: triangulating three data sources. BMC Fam Pract 2014 Apr 02;15:48 [FREE Full text] [doi: 10.1186/1471-2296-15-48] [Medline: 24690099]
42. Qualitative Research Practice A Guide for Social Science Students and Researchers. London: SAGE Publications; 2003.
43. Patel M, Chawla R, Krynicki CR, Rankin P, Upthegrove R. Health beliefs and carer burden in first episode psychosis. BMC Psychiatry 2014 Jun 09;14(1):171. [doi: 10.1186/1471-244x-14-171]
44. Fearon P, Kirkbride JB, Morgan C, Dazzan P, Morgan K, Lloyd T, et al. Incidence of schizophrenia and other psychoses in ethnic minority groups: results from the MRC AESOP Study. Psychol Med 2006 Aug 29;36(11):1541-1550. [doi: 10.1017/s0033291706008774]

Abbreviations

COPe-support: Carers fOr People with Psychosis e-support
RCT: randomized controlled trial

Edited by T Leung; submitted 21.02.21; peer-reviewed by A Rotondi, A Laine; comments to author 10.05.21; revised version received 21.05.21; accepted 30.11.21; published 02.02.22.

Please cite as:
Batchelor R, Gulshan S, Shritharan H, Williams E, Henderson C, Gillard S, Woodham LA, Cornelius V, Elkes J, Sin J
Perceived Acceptability and Experiences of a Digital Psychoeducation and Peer Support Intervention (COPe-support): Interview Study With Carers Supporting Individuals With Psychosis
J Med Internet Res 2022;24(2):e27781
URL: https://www.jmir.org/2022/2/e27781
doi:10.2196/27781
PMID:35107439

Review

Characteristics of Mobile Health Platforms for Depression and Anxiety: Content Analysis Through a Systematic Review of the Literature and Systematic Search of Two App Stores

Qiao Ying Leong[1,2], BSc; Shreya Sridhar[1]; Agata Blasiak[1,2,3,4], PhD; Xavier Tadeo[1,2], PhD; GeckHong Yeo[1,2], PhD; Alexandria Remus[1,2,3*], PhD; Dean Ho[1,2,3,4,5*], PhD

[1]N.1 Institute for Health, National University of Singapore, Singapore, Singapore

[2]The Institute for Digital Medicine (WisDM), Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

[3]Department of Biomedical Engineering, NUS Engineering, National University of Singapore, Singapore, Singapore

[4]Department of Pharmacology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

[5]Health District @ Queenstown, Singapore, Singapore

[*]these authors contributed equally

Corresponding Author:
Alexandria Remus, PhD
N.1 Institute for Health
National University of Singapore
28 Medical Drive
Singapore, 117456
Singapore
Phone: 65 86118796
Email: bieamr@nus.edu.sg

Abstract

Background: Mobile health (mHealth) platforms show promise in the management of mental health conditions such as anxiety and depression. This has resulted in an abundance of mHealth platforms available for research or commercial use.

Objective: The objective of this review is to characterize the current state of mHealth platforms designed for anxiety or depression that are available for research, commercial use, or both.

Methods: A systematic review was conducted using a two-pronged approach: searching relevant literature with prespecified search terms to identify platforms in published research and simultaneously searching 2 major app stores—Google Play Store and Apple App Store—to identify commercially available platforms. Key characteristics of the mHealth platforms were synthesized, such as platform name, targeted condition, targeted group, purpose, technology type, intervention type, commercial availability, and regulatory information.

Results: The literature and app store searches yielded 169 and 179 mHealth platforms, respectively. Most platforms developed for research purposes were designed for depression (116/169, 68.6%), whereas the app store search reported a higher number of platforms developed for anxiety (Android: 58/179, 32.4%; iOS: 27/179, 15.1%). The most common purpose of platforms in both searches was treatment (literature search: 122/169, 72.2%; app store search: 129/179, 72.1%). With regard to the types of intervention, cognitive behavioral therapy and referral to care or counseling emerged as the most popular options offered by the platforms identified in the literature and app store searches, respectively. Most platforms from both searches did not have a specific target age group. In addition, most platforms found in app stores lacked clinical and real-world evidence, and a small number of platforms found in the published research were available commercially.

Conclusions: A considerable number of mHealth platforms designed for anxiety or depression are available for research, commercial use, or both. The characteristics of these mHealth platforms greatly vary. Future efforts should focus on assessing the quality—utility, safety, and effectiveness—of the existing platforms and providing developers, from both commercial and research sectors, a reporting guideline for their platform description and a regulatory framework to facilitate the development, validation, and deployment of effective mHealth platforms.

(J Med Internet Res 2022;24(2):e27388) doi:10.2196/27388

KEYWORDS

mHealth; digital medicine; anxiety; depression; systematic review; mental health conditions; mobile phone

Introduction

Mobile health (mHealth) refers to medical, health care, and public health practices that use mobile computing or technologies such as mobile phones, wireless devices, patient monitoring devices, and web-based platforms [1-3]. With the overarching objective of improving health-related outcomes and awareness, mHealth platforms can either serve as standalone or complementary platforms and offer the potential to cater to the needs of many users, from health care professionals to patients and consumers. The existing infrastructure and technology of mobile devices, such as mobile apps and mobile sensors, make them ideal candidates for offering convenience while addressing clinical concerns and improving health-related outcomes. For example, mobile apps can be installed on a mobile device that allows for the measurement and collection of vital information such as location data and activities. These tracking data are invaluable as they may help inform users about their health status or provide necessary information to health care professionals during their decision-making process [4] to subsequently provide actionable feedback to their patients [5]. In addition, as most of the world's population already uses mobile devices, it is possible that mHealth interventions may allow greater access to care [6]. According to a digital global report published in October 2020, mobile phones and the internet had a penetration rate of 60% and 67% worldwide, respectively [7], both of which increased by a staggering 1% in just 3 months compared with July 2020 [8]. It is plausible that the COVID-19 pandemic that has driven most of the world into forceful isolation may have contributed to this increment. It is evident that mobile devices are increasingly becoming an indispensable part of our lives and will continue to be so [9]. As such, mHealth may offer a feasible approach for targeting a wide range of health conditions.

In recent years, there has been an increase in the use of mHealth platforms to intervene in the management of mental health conditions [10,11]. As a result, research is emerging on the implementation and clinical outcomes of mHealth platforms [10,11], and preliminary evidence on their efficacy is beginning to surface. Anxiety and depression, which have the highest prevalence among mental health conditions, are at the forefront of these targeted mHealth interventions [12,13]. The convenience facilitated by technology coupled with a higher quality of interventions—evidenced from randomized controlled trials (RCTs) and real-world evidence (RWE) delivered for those with anxiety or depression compared with those delivered for people with general health issues [10,11]—further propels mHealth as an increasingly appealing solution for the prevention and management of anxiety or depression. As a result, a plethora of mHealth platforms for anxiety or depression have been and continue to be developed.

This accelerated proliferation of anxiety or depression mHealth platforms that are readily available makes it difficult for stakeholders to determine not only those that may be useful but also those that may be potentially counterproductive. In addition, the resulting quantity of available platforms does not necessarily correlate with the scale of effectiveness validation. mHealth, as a whole, is still a relatively new field. Subsequently, regulatory policies and standards for the development of these platforms and the claims made by their developers have not been fully established [14,15], adding to the difficulty for stakeholders in identifying appropriate and safe mHealth apps. Accordingly, as the number of mHealth platforms for anxiety or depression continues to grow, it is important for relevant stakeholders, including users, clinicians, researchers, and developers, to understand what is currently available.

A recent review [16] has provided a comprehensive overview of commercially and academically available mental health apps offered in the Spanish language in the United States; however, a similar review of mental health mHealth apps available in the English language, to our knowledge, has not been conducted to date. Therefore, the aim of this review is to systematically identify and amalgamate the characteristics of the currently available English language mHealth platforms for anxiety or depression that have been developed for research, commercial use, or both. The outcome of this review may provide insights into considerations in the development, regulation, implementation, and adoption of future mHealth platforms for anxiety or depression.

Methods

Study Design

A cross-sectional study was performed to characterize mHealth platforms developed for depression or anxiety that were available for commercial, research, or both purposes. The mHealth platforms included apps, websites, and web-based software. A 2-pronged approach was used in this study. This included systematically searching relevant literature that described mHealth platforms as interventions developed for research and simultaneously searching 2 major app stores systematically to identify commercially available platforms. The protocol was registered on PROSPERO (International Prospective Register of Systematic Reviews; CRD42020193956) after preliminary searches were conducted. The piloting of inclusion and exclusion criteria and formal screening were conducted after the official PROSPERO registration.

Identifying Platforms Described in Research Literature

Search

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines were used and adhered to for the literature search [17]. A systematic search of the databases PubMed, Embase, CINAHL, and PsycINFO was completed on June 23, 2020, from database inception dates. The search string was developed using the population, intervention, comparison, outcome strategy [18] using keywords and search terms combined with Boolean operators (Textbox 1). The search terms were carefully selected so that the homogeneous terms could be used in both the literature and app

store searches. The full search strings per database can be found in Multimedia Appendix 1. All searches were performed with no restriction on the publication period or year of release. There were no age or location restrictions on the target population.

Textbox 1. Keywords and search terms used for the formulation of the search string in the literature search.

Mental health condition

- *Mental health, mindfulness, anxiety, depression*

Platform

- *mHealth, mobile health, mobile device, mobile application, digital therapeutics, digital intervention, ehealth, smartphone, mobile phone, text message, web based, web application*

Selection

After duplicate articles were removed, 2 reviewers (SS and QYL) independently applied the eligibility criteria to titles and abstracts and subsequently to full-text articles using the Covidence software (Veritas Health Innovation) [19]. A third reviewer (AR) resolved any discrepancies. Articles were included if they were published in English, were full-text original journal articles or conference papers, and provided at least one characteristic of interest about an mHealth platform for anxiety or depression.

Articles were excluded if they were reviews (but reference lists were cross-checked for included studies), had no characteristic description of the mHealth platform, did not have the name of the mHealth platform, were not designed specifically for anxiety or depression, described mHealth platforms that targeted physical conditions only, or targeted secondary mental health problems related to a physical condition, health behaviors (eg, smoking cessation, risky drinking, obesity, and exercise), neurodegenerative disorders (eg, dementia), chronic pain, a specific phobia, posttraumatic stress disorder, addictions, subclinical symptoms, or a severe mental illness.

As the primary focus of the review was to characterize mHealth platforms, a formal assessment of the study quality and risk of bias was not performed.

Identifying Platforms Available Commercially

Search

Building on the methodologies by Bender et al [20] and Giunti et al [21], we systematically completed our search of all available platforms from 2 major app stores accessible from Singapore—Google Play Store and Apple App Store—on June 19, 2020. In-depth searches across the Apple App Store were included to identify platforms compatible with iOS or macOS devices. The search terms used were *anxiety*, *depression*, *mental health*, and *mindfulness* (Textbox 1), with no restriction on the app category.

Selection

Apps were initially included if the title or description of the app included the terms anxiety or depression. After the removal of duplicates, 2 reviewers (SS and QYL) independently assessed the eligibility of the app using the a priori determined criteria. Any discrepancies were addressed by a third independent reviewer (AR).

The included mHealth platforms comprised apps targeting mental health, specifically anxiety, depression, or both. Apps were excluded if they were no longer available, not available in English, or provided app descriptions that lacked the characteristics of interest.

Data Coding and Extraction

Data Coding

An mHealth platform characteristic coding scheme was developed based on an internally identified mHealth platform's characteristics of interest. If the mHealth platform characteristics were not clearly described or discrepancies were found during the classification of the mHealth platforms, discussions were held until consensus was reached. The final coding scheme and outcomes are detailed in Table 1.

Table 1. The mobile health (mHealth) characteristic coding scheme of the characteristics of interest from the respective searches and the description of each item.

Characteristic	Description	Literature search	App store search
Platform name	The name of the platform that was developed for research, available for commercial use, or both	✓	✓
Targeted condition	The condition the platform was designed for: anxiety, depression, or both	✓	✓
Targeted group	The intended users the platform was designed for	✓	✓
Commercial availability	The commercial availability of the platform in either the Apple App Store, Google Play Store, or both	✓	✓
Purpose of platform	The intended use that the platform was designed for. For example, diagnostics, monitoring, prevention, treatment, education, and support	✓	✓
Type of technology	The mobile modality and operating system in which the mHealth platform was rolled out	✓	✓
Type of intervention	Applicable only for platforms that offer treatment	✓	✓
App store categorization	App store category assigned to the platform by the development team that best describes its main function or subject matter		✓
Additional mHealth characteristics	Additional characteristics reported such as language, ratings, and cost of the platform	✓	✓
Clinical evidence and regulatory information	A secondary search was performed to obtain additional information of the identified platforms which included the availability of clinical evidence in the form of randomized control trials, real-world evidence, and regulatory approval	✓	✓
Other studies that reported the platform	Studies cited in the platform's article regarding the development, validation, or further evaluation of the platform	✓	
Cited literature	References to academic publications provided in the platform's app store description		✓

Data Extraction

The mHealth platforms and their characteristics of interest from both searches were extracted and recorded in a custom-developed Microsoft Excel template. The collected and assessed data included characteristic information of the mHealth platforms, which were based on the description in the app stores, provided by 2 reviewers (SS and QYL), and any discrepancies were resolved by a third reviewer (AR).

Statistical Analysis

Descriptive statistical analyses were performed for all the variables. The categorical variables were presented as absolutes and relative frequencies. Statistical analyses were performed using Microsoft Excel (version 16.43).

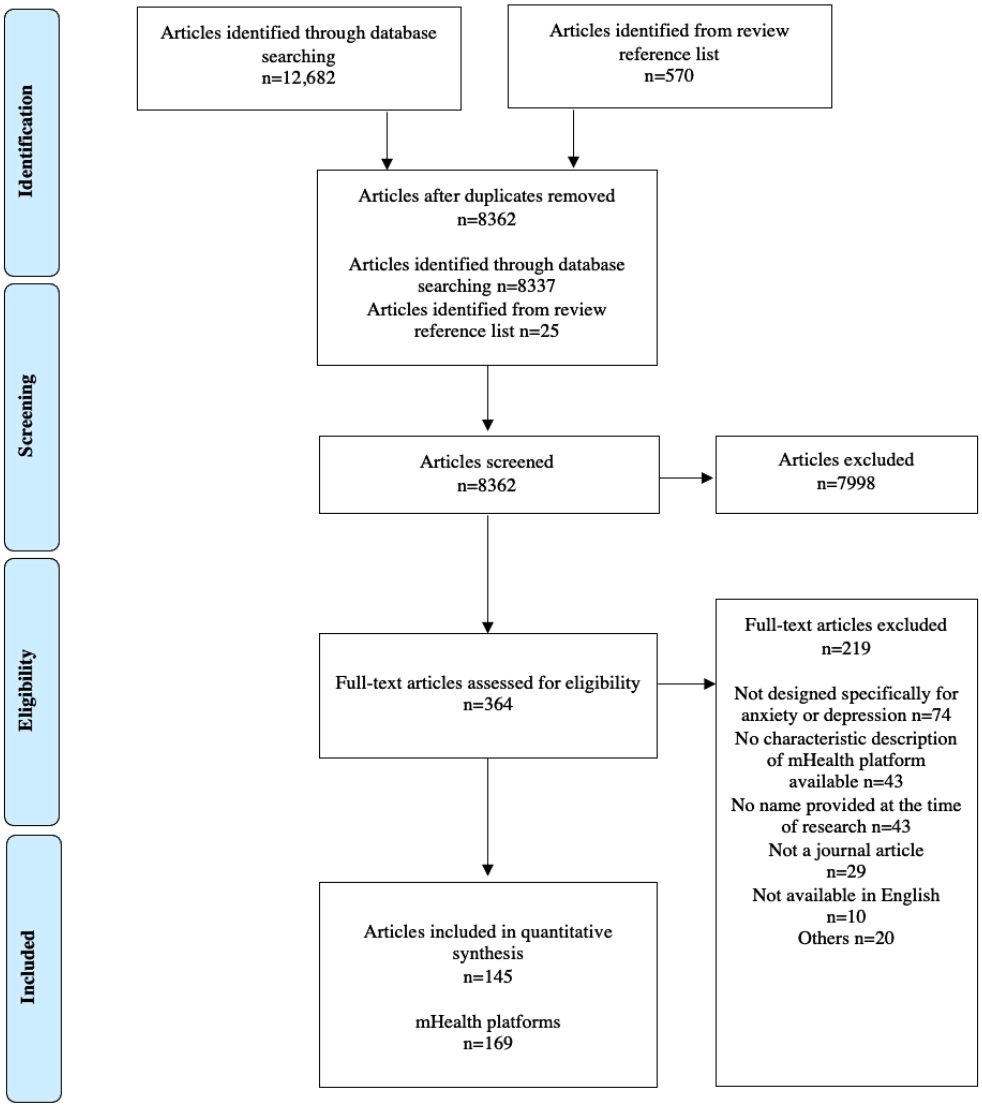
Results

Platforms Available in Literature Search

Selection

As depicted in Figure 1, after duplicate removal and review of reference list cross-checking, 8362 articles were initially screened. This resulted in the identification of 169 mHealth platforms described in 1.7% (145/8362) of articles (Multimedia Appendix 2 [22-166]). Some articles reported ≥1 platform, whereas some platforms were described in ≥1 article. Examples of platforms found across multiple studies were Deprexis (6/169, 3.6%), MoodGYM (6/169, 3.6%), Happy@Work (4/169, 2.4%), myCompass (4/169, 2.4%), and Partners in Parenting (4/169, 2.4%).

Figure 1. PRISMA (Preferred Reporting Item for Systematic Reviews and Meta-Analyses) flow diagram of search strategy and the finalized number of articles and resulting mHealth platforms included for this review. mHealth: mobile health.



Commercial Availability of Apps From the Literature Search

Of the 169 apps identified in the literature search, 8 (4.7%) were available in the app stores. All 8 apps were found in the Google Play Store, whereas a subset (6/8, 75%) was commercially available in the Apple App Store. No platform was reported as a medical device. However, of the 169 platforms, 2 (1.2%)—myStrength and HelpID—were reported as being used in health care settings. myStrength was used as a psychological intervention and had affiliations with health care providers, whereas HelpID was included in health care plans and insurance.

Targeted Condition

Of the 169 platforms reported in the articles, 116 (68.6%) targeted depression, 22 (13%) targeted anxiety, and 31 (18.3%) targeted both (Table 2). Approximately 1.8% (3/169) platforms—Situman, TeleCoach, and LifeRhythm—acted as complementary apps to other platforms targeting depression. The roles played by these apps varied, where Situman provided situational awareness to the Moodbuster platform, TeleCoach involved weekly calls to help with adherence to the MoodManager intervention, and LifeRhythm helped in data collection for the DepWatch system.

Table 2. Summary of the mobile health (mHealth) platforms characteristics from the literature search (N=169).

mHealth platform characteristics	Platforms, n (%)
Targeted condition	
Anxiety	22 (13)
Depression	116 (68.6)
Both	31 (18.3)
Target group	
Children	3 (1.8)
Young people	15 (8)
University students	3 (1.8)
Adults	8 (4.7)
Older adults	1 (0.6)
Employees	3 (1.8)
Migrants	1 (0.6)
Patient or clinician	2 (1.2)
Not specified	136 (79.9)
Purposes	
1 only	90 (53.3)
2	71 (42)
≥3	8 (4.7)
Number of languages	
1	2 (1.2)
2	5 (3)
≥3	4 (2.4)
Not specified	158 (93.5)

Targeted Group

Of the 169 platforms reported in the articles, 135 (79.9%) were not developed for a specific target group. The remaining 20.1% (34/169) targeted 8 different groups: children, young people, university students, adults, older adults, employees, migrants, and patient or clinician communication. Of the 169 apps, 2 (1.2%)—BiP Anxiety and Supporting Our Valued Adolescents—involved ≥1 target group and required the engagement of parents along with their children and adolescents, respectively.

Purpose

Approximately half of the platforms were designed with ≥1 purpose (Table 2). Treatment was the most common purpose, with 72.2% (122/169) of the platforms including it (Table 3).

Table 3. The primary purpose of mobile health platforms for anxiety or depression identified in the literature search (N=169).

Purpose	Description	Example	Platforms, n (%)
Education	• Provide mental health and medication information • Teach a variety of skills, including self-management and problem solving • Impart lessons through modules	Text-based information, audio or visual materials, and blogs	45 (26.6)
Diagnostic	• Identify depression and anxiety symptoms • Serve as a screening tool	Psychological questionnaires before use or over time	12 (7.1)
Monitoring	• Track self-reported symptoms and medical adherence • Collect relevant sensor data such as location and mobility information	Ecological momentary assessment through a mobile app, prompts to rate mood or complete short questionnaires, and data collection through wearable devices and mobile phones	38 (22.5)
Treatment	• Deliver psychological interventions through modules • Teach a variety of skills such as positive emotion skills, self-management skills, and guided meditation • Learn about coping strategies	Cognitive behavioral therapy, problem-solving therapy, gamification, lessons by topic, and homework	122 (72.2)
Prevention	• Similar to Treatment but targets individuals who are at risk of anxiety or depression	Text-based information and audio or visual, synchronous, or offline asynchronous activity	29 (16.2)
Support	• Offer a decision support system to clinicians • Manage care activities • Provide access to peer-to-peer support or health care professionals for support	Messaging service, discussion forum, web-based interaction in self-help groups, and web-based directory of bodies offering support	11 (6.5)

Type of Intervention

In platforms designed for treatment purposes, the incorporation of ≥2 intervention types was common (20/169, 11.8%). The most common types of intervention were cognitive behavioral therapy (CBT; 50/169, 29.6%), problem-solving therapy (11/169, 6.5%), and psychoeducation (9/169, 5.3%). An average of 8 (SD 5.04) modules was offered by 41.4% (70/169) of platforms that provided module-based sessions. The common frequency of intervention was weekly (32/169, 18.9%), followed by daily (8/169, 4.7%) on the (42/169, 24.9%) platforms that reported this characteristic. Of note, some (5/169, 3%) platforms incorporated some form of in-app human interaction, where the interaction personnel included medical physicians (1/5, 20%; Ascend), psychologists (1/5, 20%; HelpID), therapists (1/5, 20%; SmartCAT), trained health care professionals (1/5, 20%; PratenOnline), Master's degree–level students in clinical psychology (2/5, 40%; Happy@Work and Ascend), and occupational social workers (1/5, 20%; Happy@Work).

Type of Technology

Various types of technology have been deployed to fulfill the purposes of the mHealth platforms. Some (22/169, 13%) platforms used ≥1type of technology, with top choices including web-based interventions (90/169, 53.3%) and mobile apps (57/169, 33.7%; Table 4).

Table 4. Types of technology used by the platforms made for research purposes (N=169).

Technology type	Platforms, n (%)
Web-based platforms	90 (53.3)
Mobile apps	57 (33.7)
Text messaging	12 (7.1)
Others	6 (3.6)
Not specified	25 (14.8)

Additional mHealth Characteristics

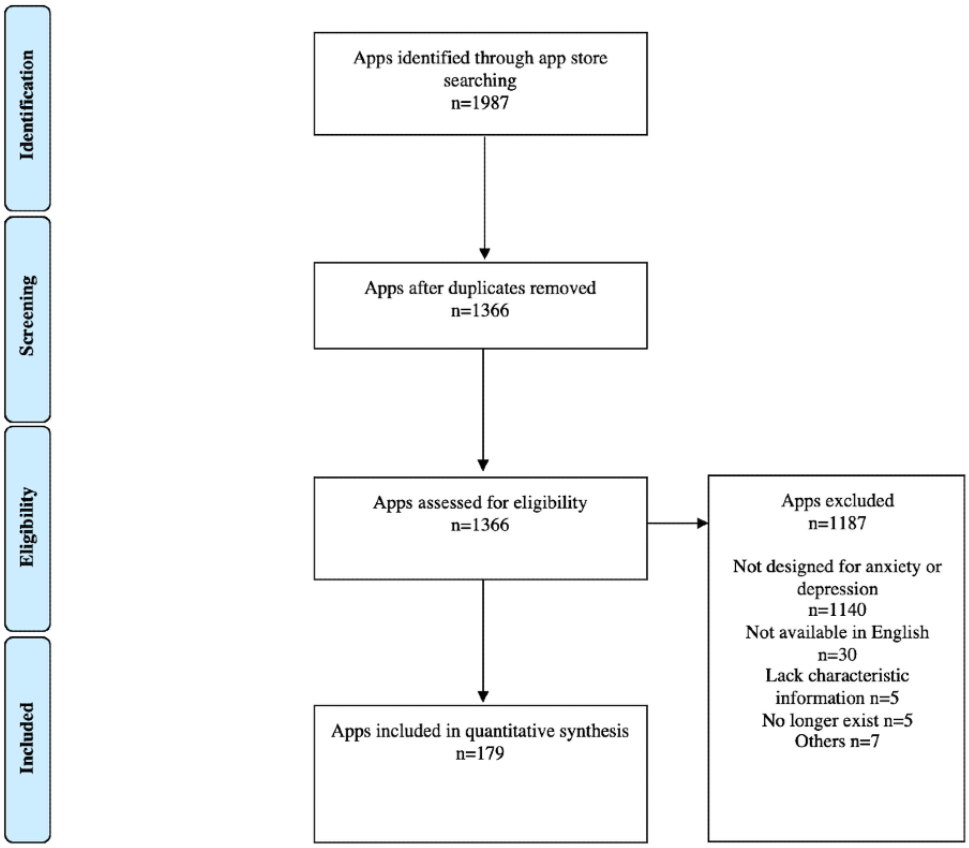
Of the 169 platforms, 9 (5.3%) were available in ≥2 languages (Table 2). Of these, the most common languages besides English were German (4/9, 44%) and Spanish (4/9, 44%).

App Store Search

Selection

The search terms yielded 179 mHealth platforms from the selection of apps in the Singapore region's Google Play Store and Apple App Store (Figure 2; Multimedia Appendix 3).

Figure 2. PRISMA (Preferred Reporting Item for Systematic Reviews and Meta-Analyses) flow diagram of search strategy and the finalized number of platforms included in the app store search for this review.



Commercial Availability

More apps were available in the Google Play Store (140/179, 78.2%) than in the Apple App Store (61/179, 34.1%; Table 5). Approximately 12.3% (22/179) apps were available in both app stores.

Table 5. Summary of characteristics of commercially available mHealth platforms (N=179).

Characteristics	Google Play Store (n=140), n (%)	Apple App Store (n=61), n (%)	Both[a] (n=22), n (%)	All, (n=179), n (%)
Targeted condition				
Anxiety	58 (41.4)	27 (44.3)	12 (54.6)	73 (40.8)
Depression	39 (27.9)	13 (21.3)	2 (9.1)	50 (27.9)
Both	43 (30.7)	21 (34.4)	8 (36.4)	56 (31.3)
Purposes				
1	77 (55)	32 (52.5)	7 (31.8)	102 (57)
2	45 (32.1)	16 (26.2)	7 (31.8)	54 (30.2)
≥3	18 (12.9)	13 (21.3)	8 (36.4)	23 (12.9)
Frequency of intervention				
Daily	4 (2.8)	3 (4.9)	0 (0)	7 (3.9)
Weekly	1 (0.7)	2 (3.3)	1 (4.6)	2 (1.1)
Not specified	135 (96.4)	56 (91.8)	21 (95.5)	170 (95)
Number of downloads				
<5000	67 (47.9)	4 (6.6)	4 (18.2)	67 (37.4)
5000-10,000	12 (8.6)	1 (1.6)	1 (4.6)	12 (6.7)
10,000-50,000	26 (18.6)	3 (4.9)	3 (13.6)	26 (14.5)
50,000-100,000	8 (5.7)	0 (0)	0 (0)	8 (4.5)
100,000-500,000	14 (10)	5 (8.2)	5 (22.7)	14 (7.8)
500,000-1,000,000	3 (2.2)	1 (1.6)	1 (4.6)	3 (1.7)
>1,000,000	6 (4.3)	4 (6.6)	4 (18.2)	6 (3.4)
Not available	4 (2.9)	43 (70.5)	4 (18.2)	43 (24)
Ratings[b]				
1 star	1 (0.7)	1 (1.6)	0 (0)	2 (1.1)
2 stars	0 (0)	1 (1.6)	0 (0)	1 (0.6)
3 stars	7 (5)	4 (6.6)	0 (0)	10 (5.6)
4 stars	58 (41.4)	21 (34.4)	10 (45.6)	70 (39.1)
5 stars	43 (30.7)	33 (54.1)	12 (54.6)	64 (35.8)
Not specified	31 (22.1)	1 (1.6)	0 (0)	32 (17.9)
Cost				
Free	85 (60.7)	19 (31.2)	4 (18.2)	100 (55.9)
Free and paid versions available	39 (27.9)	30 (49.2)	13 (59.1)	56 (31.3)
Payment required	16 (11.4)	12 (19.7)	5 (22.7)	23 (12.8)

[a]Both refer to platforms found in both Google Play Store and the Apple App Store.

[b]Ratings were rounded to the nearest integer.

Targeted Condition

The mHealth apps identified in both app stores were designed predominantly for anxiety (73/179, 40.8%). Approximately 30.7% (43/140) of the apps in the Google Play Store and 34% (21/61) of the apps in the Apple App Store claimed to address both anxiety and depression (Table 5).

Targeted Group

All platforms provided suitable age groups. One of the platforms, the Geriatric Depression Scale 2.0, detailed a specified target group, that is, older adults.

Purpose

Most apps found in both stores offered 1 purpose (Android devices: 77/179, 43%; iOS devices: 32/179, 17.9%; Table 5). Among the identified purposes available through the apps,

treatment was the most common purpose in both app stores, with 72.1% (129/179) of the apps including it (Table 6).

Table 6. The primary purpose mobile health platforms for anxiety or depression identified from the Google Play Store and Apple App Store (N=179).

Purpose	Description	Examples	Apps, n (%)
Education	• Provide information and communication of self-monitoring procedures • Dissemination of information by health care professionals	Text-based information, video materials, and SMS text messaging	50 (27.9)
Diagnostic	• Interpret symptoms • Identify possible mental health conditions	Self-administered questionnaires and psychological tests	34 (19)
Treatment	• Offer treatment tips, recommendations, and advice • Provide some form of intervention	SMS text messaging, real-time videoconferencing, and text-based treatment suggestions	129 (72.1)
Monitoring	• Tracking of symptoms • Prompts and reminders for tasks and appointments	In-app notifications, self-administered questionnaires, and journaling	26 (14.5)
Support	• Provide access to some form of support • Features option to call a friend, a family member, health care professionals, or a nearby help center directly from the platform	Forums, chatbots, and SMS text messages	34 (19)
Prevention	• Claim to prepare for and prevent anxiety-related or depressive episodes • Offer modules or activities that serve as preventive measures	Text-based information, audio, and activities	8 (4.5)

Type of Intervention

A total of 22 types of interventions were identified from the 72.1% (129/179) of apps that offered some form of treatment. Of these, referral to care or counseling (36/129, 27.9%), problem solving (27/129, 20.9%), and CBT (22/129, 17.1%) were the most common types of interventions available. Apart from encouraging to seek or directly offering specialist support, some apps provided support from the community, for example, access to forums with like-minded users (16/179, 8.9%). A subset of the apps that offered referral (14/179, 7.8%), counseling (4/179, 2.2%), or forum services (2/179, 1.1%) was subscription-based or required payment to gain access.

Type of Technology

Of the 179 platforms, all the platforms were available for mobile devices, except for 1 (0.6%; At Ease Anxiety and Worry Relief), which was accessible from both iOS and macOS devices. Of these 179 platforms, 140 (78.2%) were available on the Android operating system, 61 (34.1%) were available on iOS, and 22 (12.3%) were available on both operating systems. Other notable forms of technology that complemented or were incorporated in the platforms included wearable devices, such as headsets or smartwatches (1/179, 0.6%), and artificial intelligence–backed counseling sessions (3/179, 1.7%).

App Store Categorization

The app categorization and classification differed between the 2 stores. In Google Play Store, mHealth apps were categorized into 14 categories (Health and Fitness, Medical, Lifestyle, Books and Reference, Education, Trivia, Casual, Simulation, Adventure, Puzzle, Music and Audio, Social, Personalization, and Productivity), whereas Apple App Store mHealth apps were categorized into 5 categories (Health and Fitness, Medical, Lifestyle, Games, and Social Networking). The Health and Fitness category was the most common classification of mHealth platforms available on both the Google Play Store (81/179, 45.3%) and Apple App Store (39/179, 21.8%). The Medical category was the second most common classification available on both the Google Play Store (27/140, 19.3%) and Apple App Store (18/61, 29.5%).

Additional mHealth Characteristics

Of the 179 mHealth platforms, 100 (55.9%) were made freely available, with 56 (31.3%) providing both free and paid versions and 23 (12.8%) offering only paid versions. The costs of platforms that required payments ranged from SGD 1.21 (US $0.90) per installation to SGD 399.99 (US $296.20) for lifetime use.

Clinical Evidence and Regulatory Information for Literature and App Store mHealth Platforms

RCTs are clinical evidence and a prerequisite for regulatory approval. However, not all platforms had RCTs. Additional regulatory information—RWE and regulatory approval—was searched. In the secondary search mentioned in Table 1, a subset of apps had a credible source that offered some form of empirical evidence. In the app store searches, 3.4% (6/179) of the apps included scientific evidence claims, whereas 2.8% (5/179) of the apps were affiliated with health care experts and reputable institutions.

None of the apps in both literature and app store searches were reported as medical devices. In the literature search, 44.4% (75/169) and 45% (76/169) of articles presented evidence from

RCTs and RCT protocols (Table 7), respectively, whereas the remaining 10.7% (18/169) did not provide evidence from RCTs. Further analysis of the clinical and regulatory evidence—through the search for additional RCTs, RWE, and regulatory approvals—demonstrated that 41.4% (70/169) of platforms presented evidence from RCTs, 0.6% (1/169; Wysa) presented preliminary RWE, 1.2% (2/169; myStrength and Moodgym) presented evidence from RCTs and RWE, and 0.6% (1/169; Deprexis) presented evidence from RCTs and RWE and had obtained regulatory approval.

From the app store search, a subset of apps was pinpointed to be based on substantial validation studies to assess effectiveness. In particular, RWE can serve as a key gateway and vitally important indicator for supporting health care decisions, as well as impact and value toward patient care and the broader health care ecosystem. For example, 1.1% (2/179) of the apps provided evidence from each of the following sources: RCTs (Sanvello and Moodmission); RWE (Talkspace and Wysa); RCTs, RWE, and regulatory approval when applicable (Flow and Sooma).

Table 7. Availability of clinical-grade evidence and regulatory oversight in mobile health platforms for anxiety or depression identified from the literature and app store searches.

Evidence	Literature search (N=169), n (%)	App stores search (N=179), n (%)
RCTs[a] protocol	75 (44.4)	—[b]
RCTs	76 (45)	2 (1.1)
RWE[c]	—[b]	2 (1.1)
RCT+RWE	—[b]	0 (0)
RCT+RWE + regulatory approval	—[b]	2 (1.1)
None	18 (10.7)	173 (96.7)

[a]RCT: randomized controlled trial.

[b]No reported clinical evidence or additional regulatory information found.

[c]RWE: real-world evidence.

Discussion

Principal Findings

This review provides an in-depth analysis of the current mHealth platforms developed for anxiety or depression available for research and commercial purposes. Through a novel combined approach—a 2-pronged systematic search across the research literature and marketplace—we identified 169 and 179 mHealth platforms, respectively. Most platforms developed for research purposes were designed for depression (116/169, 68.6%), whereas the platforms made available in the app stores were mainly developed for anxiety (Android: 58/179, 32.4%; iOS: 27/179, 15.1%). We identified that mHealth characteristics such as platform name, aspects of mental health addressed, purpose, type of intervention, type of technology, and level of evidence varied largely across these platforms. In addition, we highlighted a significant lack of adoption of the regulatory framework and standards governing these mHealth platforms. An overview of the existing platforms for anxiety and depression available for research, commercial use, or both, as reported in this review, allowed us to identify gaps in the literature and marketplace and subsequently provide guidance for recommendations for future frameworks of mHealth development and regulations.

Targeted Group, Purpose, and Intervention Type of Available mHealth Platforms for Anxiety and Depression

Among the abundance of mHealth options currently available for anxiety or depression, relatively few have been developed for youths and emerging adults [167,168]. This is surprising, as these age groups are among the most prevalent users of mobile technologies as well as the groups that are most in need of mental health interventions. In particular, typical engagement with standard care for mental health has low appeal to young people and can be antagonistic to their developmental needs [168]. The low motivation inherent in emotional disorders; the need for agency, anonymity, and privacy; the greater sensitivity to stigma and discrimination associated with mental health; and the development of self-monitoring among young people can be addressed by mHealth approaches that provide personalized and timely treatment [169,170]. Future developers of mHealth platforms for anxiety or depression should consider developing them with this targeted group in mind.

The main purpose of most of the mHealth platforms for anxiety or depression identified in this review focused on treatment. The treatment used features such as module-based sessions and the teaching of coping and management strategies, which were based on the principles and guidelines of CBT, problem-solving therapy, and psychoeducation. Previous research has revealed 3 primary areas for enhancing the effectiveness of the treatment function of mHealth [167]. First, mHealth needs to be complemented by human interactions, especially with a medical physician or health care worker, to increase compliance, adherence, dropout [2,171,172], and efficacy in treatment [173-175]. However, findings from this review highlighted that few mHealth platforms included human interactions. To be effective, the digital framework for human interaction needs to consider individuals' desire for anonymity and privacy. Second, initial engagement and sustained use of mHealth platforms for young people require some form of social support from peers, schools, or mental health professionals. Although the role of social support and the associated outcomes are similar to that

found with the standard of care, the distinct feature of digital social support that mHealth platforms provide is anonymity and accessibility. The opportunity to share similar lived experiences of anxiety or depression with one's peers and being accountable to an authority (eg, health care professional) for the purpose of monitoring or assessment of their symptoms has the potential to increase their engagement and adherence to the mHealth platforms. Notwithstanding the support function of mHealth platforms, few platforms included this in their design feature. Finally, an mHealth platform needs to be easy to navigate, relatable, engaging, and esthetically appealing in providing education for mental health users—another common function of mHealth platforms. Previous studies have found that the boring and repetitive nature of modules and lessons presents a barrier to greater engagement with the education feature of an mHealth platform [10].

Type of Technology

mHealth can be deployed via one or a combination of mobile technologies. An increasing trend, which is also observed in health care, is to reach the individuals by communication technologies that are already in use for the seamless incorporation of the intervention into their daily life. Our literature search reported web-based platforms (90/169, 53.3%), mobile apps (57/169, 33.7%), and messaging (12/169, 7.1%) as the most common modalities deployed, which is consistent with findings in other studies [176]. In our literature search, 13% (22/169) of the platforms used ≥1 type of technology.

The app stores search identified discrepancies in the number of apps available in each app store. The higher number of apps in the Google Play Store (140/179, 78.2%) than in the Apple App Store (61/179, 34.1%) can be attributed to a preference for the Android operating system over iOS in Asia (82.7% vs 16.5%) and Singapore (65.2% vs 33.9%), as recorded in June 2020 [177]. Although the infrequent download count data from iOS disables the comparison in this category in our review, the literature suggests that in the United States, the iOS preference is reversed (41.7% vs 58.2% for Android and iOS, respectively) [177], and iOS apps show a higher number of downloads of behavioral health mobile apps [178]. Interestingly, a recent study collecting location data has shown that the platform type had a measurable effect on the retention of passive data collection because of the increased battery consumption of the app for the iOS system. In addition, the operating systems differ in their battery-saving modes and memory space–saving strategies, for example, app offloading to a cloud, which may have further implications. In the same study, no differences were detected in active data collection retention, suggesting that the operating system had no effect on active user interactions with the mHealth platform [179].

The deployment technology type and its specific technical considerations (eg, operating system) can *make or break* an mHealth platform. A digital intervention design should carefully consider the context of how each technology is already being used, as well as its technical and behavioral limitations. An mHealth platform can be deployed either via one technology type only or in a hybrid model to reap the benefits of each technology type to maximize usability with sustained functionality toward the highest efficiency of the intervention.

Commercial Versus Academic Motivation Considerations

Of note, only 4.7% (8/169) of the mobile apps identified in the literature search were available in app stores. This can be potentially related to the geographical constraints of the app store search compared with the unconstrained geography of the literature search. In addition, the literature searches covered articles from inception to 2020, whereas the app store search included only the apps present in the marketplace in June 2020. It is not unusual that after the gathering of evidence, the apps become unavailable in the marketplace [178,180,181]. The lack of commercial translation of research published in literature can be attributed to multiple factors: funding timelines (eg, relatively short grant duration and high staff rotation may not enable long term app management) and downstream research objectives (basic or fundamental research, tool development, and proof-of-concept to support potential future commercial and large-scale deployment) [182].

From the perspective of the marketplace, 1.7% (3/179) of apps in the app store search were described in a peer-reviewed article identified in our search. This can be potentially attributed to a business strategy that favors *quick-in* and *quick-out* rather than laborious and costly app evaluation without prevailing guidelines or incentives—monetary, regulatory, or otherwise—in doing so. Although consumers make the initial decision of which app to download based on easy-to-judge attributes such as graphics, price, and rating [178,183], it is the credibility and trustworthiness that drive engagement with mental health apps [184-186]. With this realization, it is important for commercial app developers to generate scientific evidence that validates their apps to substantiate the scientific descriptions, which will collectively enhance the app's appeal. In a study of the app descriptions of the top-ranked apps, it was previously found that apps use scientific jargon and are not supported by principles backed by peer review and validation studies, or they do not align with established findings in the scientific literature [187]. Instead of objectively informing about the evidence, the function of the app description is to appeal to consumers and drive the number of app downloads, independent of the platform's objective quality [187]. Accordingly, app user ratings do not indicate clinical utility or quality [188,189].

Regulatory Information: Scientific Evidence and Regulatory Status of mHealth Platforms for Research and Commercial Use

Clinical evidence, in the form of RCTs, reported in the publications found in the literature search (151/169, 89.4%) vastly exceeded that for apps found in the app store (2/179, 1.1%). Regulatory approvals (0.6% vs 1.1%) and RWE (0.6% vs 1.1%) were more balanced between platforms from the literature and app stores. The relatively high frequency of RCTs but low numbers in RWE and regulatory approvals for platforms featured in the literature search seem to indicate a pre-eminently investigative intent of those platforms. In addition, 33.7% (57/169) of the platforms were available as mobile apps, further hinting at treatment commercialization as a secondary

possibility. Maintenance and update of the website platforms were, in some cases, not frequent, indicating ad hoc use.

A large proportion of the mHealth platforms from the app store were unregulated, clinically untested, and unevaluated for real-world effectiveness (173/179, 96.7%), indicating a potential need for business models that support investment in robust clinical or real-world validation. Importantly, although 1.7% (3/179) of the apps found in the app store search in this review could claim to have scientific evidence or regulatory approval, 15.1% (27/179) and 10.1% (18/179) of the mHealth platforms for anxiety or depression extracted from the Google Play Store and Apple App Store, respectively, were included in the Medical category. Under this category (excluding telehealth), 8.4% (15/179) of the platforms affirmed that health professionals were involved in the design of the apps. From the app stores, 8.9% (16/179) of the platforms included a disclaimer notifying that the app was not intended to be a replacement for treatment or any sort of medical intervention. Another 12.3% (22/179) of the apps mentioned information sources used to design the app, such as validated tests (eg, Patient Health Questionnaire-9 for the assessment of depression and Generalized Anxiety Disorder-7 for the assessment of anxiety), international organization guidelines (eg, World Health Organization), validated therapies (eg, CBT), or peer-reviewed articles. These numbers reflect the need to develop more actionable and meaningful app store categorizations for medical apps. In addition, increased app store categorization oversight and strategies to enable or incentivize owners to implement changes, as well as clearer accountability guidelines pertaining to store content quality, may be needed. For example, oversight could be extended to app store owners to assist in enhancing specificity regarding the medical classification of apps. In addition, greater user education in the selection of apps may be a downstream consideration for health authorities. Consumers may need to be able to discern real-value apps by analyzing their claims, the involvement of health professionals, regulatory status, and scientific evidence. Apps that have not been supported by established validation guidelines adversely affect the field and users in multiple ways. For example, they may add a burden to already strained health care systems. In addition, they may create opportunity costs when the technological promise is not realized, resulting in mHealth fatigue of the clinicians, patients, and consumers, thus delaying the implementation of evidence-based, efficient mHealth platforms.

Overall, a need for enhanced regulation, especially in the Medical category of the stores, indicates that authorities may be able to positively affect the mHealth landscape toward requirements that include threshold validation and other applicable compliance guidelines. Several attempts have been made in that direction from various government organizations and professional societies. For example, the Institute of Electrical and Electronics Engineers—an engineering professional society—developed the Institute of Electrical and Electronics Engineers 1752 standards in 2017, which define specifications for standardized representations for mHealth data and metadata for cardiovascular, respiratory, and metabolic measures [190]. Health Level Seven International—an international mHealth professional workgroup—released the Consumer Mobile Health Application Functional Framework in 2018 with the primary goal of providing a standard whereby a mobile app's characteristics can be assessed [191]. In 2020, the American Psychiatric Association—a professional organization of psychiatrists—developed the American Psychiatric Association App Advisor, an initiative to identify an app evaluation model that was specific to mental health to be used by clinicians and their patients [192]. Most recently, in February 2021, a joint unit formed by the National Health Service England and the Department of Health and Social Care—released the Digital Technology Assessment Criteria with the aim to set the expectation of digital health technologies for potential developers [193]. In addition, other digital health and medicine professional societies, such as the Digital Medicine Society and Digital Therapeutics Alliance, also have initiatives to develop frameworks, standards, and resources for the different stakeholders of digital health interventions [194,195]. Furthermore, the International Organization for Standardization Technical Specification, in collaboration with the European Committee for Standardization, is currently developing a health app quality label *ISO/TS 82304-2 Health software–Part 2: Health and wellness apps–Quality and reliability* [196,197]. Although the continued development of these regulatory frameworks is critical, to have a real-world impact on the quality of mHealth platforms and apps, they must also be widely adopted by developers and requested by regulators, marketplace owners, and target populations alike. This will ultimately result in the development of high-quality mHealth platforms with stringent safety and security measures required in health care.

Methodological Limitations and Further Considerations

Although we took a novel approach of combining both a systematic literature search and app store search, some mHealth platforms may have been excluded. For example, although we had no geographical filters in our literature search, our commercially available app search was based on regional availability in Singapore's Google Play Store and Apple App Store. Although this may be a region-specific search, this review also provides new insight into the field as it serves as an early resource detailing available mHealth apps and platforms in South-East Asia. As Singapore is a globally recognized regional information communications technology hub with a well-established and efficient health care innovation ecosystem, this review provides the basis for a review of commercially available mHealth platforms available in surrounding countries and worldwide, alongside actionable recommendations. In addition, we restricted our inclusion criteria to mHealth platforms in English. It is possible that some apps may be published exclusively in other countries' stores (eg, United States, United Kingdom, or Canadian stores) and in other languages that were ultimately excluded from our search and this review. Furthermore, in the academic setting, the goal of the research team may be to assess the efficacy of the mHealth intervention they developed. After proven efficacious, commercialization may be the next step in the pipeline, in which case a name may be given to the platform. As such, this review does not include nameless mHealth platforms that were developed solely for research purposes. This decision was made

so that we could identify the app crossover between the 2 searches. In addition, we used exclusion criteria such as mHealth platforms developed for specific target populations (eg, comorbidities), secondary mental health problems related to a physical condition (eg, pregnancy), and platforms targeting health behaviors (eg, anxiety for smoking cessation) and severe mental illnesses. These criteria narrowed down the included platforms considerably. Overall, the possibility of mHealth platform exclusion may have introduced bias into our analysis. However, by including a comprehensive search of both the literature and multiple app stores, we aimed to minimize the possibility of mHealth platform exclusion in this review. Finally, as previously mentioned, we carefully selected search terms that could be used in both the literature and app store searches to ensure that these searches could be conducted as homogeneously as possible despite the different methodological requirements for each. As a result, some common key terms (eg, treatment and prevention) and phrases for the clinical conditions (eg, mood, sadness, affective disorder, phobias, and CBT) that would be included in a traditional literature database search were excluded. Therefore, it is plausible that some mHealth platforms described in the literature may have been excluded from this review. However, this review aimed to provide an overview of current mHealth platforms developed specifically for general nonspecific anxiety or depression; thus, we do not expect many potentially excluded platforms from this review. If more specific anxiety and depression conditions such as mood disorders or phobias are of interest, the keywords should be developed accordingly.

Furthermore, this review does not include a quality assessment of the mHealth platforms. We intended to use the Mobile App Rating Scale [198], as indicated in our PROSPERO registration (CRD42020193956). However, it was beyond the scope of this project to download all the included mHealth platforms, many of which required payment for installation or were not accessible (ie, from the literature). Therefore, we used the information detailed in the descriptions provided by the app stores or in the literature. As a result, some data were not available, and we could not avail of the Mobile App Rating Scale for the intended purpose. We attempted to use an alternative mHealth quality assessment tool; however, we did not find one that would be suitable for these purposes. Nouri et al [199] has previously identified substantial heterogeneity in the assessment criteria for mHealth apps in different studies. As valid and reliable frameworks to assess the quality of mHealth platforms are developed, future reviews, such as ours, should incorporate them.

Future Directions: Digital Therapeutics as the Next Step for mHealth Solutions for Anxiety and Depression

It is estimated that 264 million people have depression and 284 million people have anxiety worldwide [200]. The COVID-19 pandemic may exacerbate these numbers [201]. Although effective treatments exist, many health systems worldwide are underresourced and struggling to respond to the burden of anxiety, depression, and other mental health disorders effectively and efficiently [201]. As a result, mHealth platforms continue to gain popularity and are emerging as a feasible option for improving the quality of and overcoming access barriers to mental health support. However, the aforementioned concerns of lack of regulation and low levels of evidence are limiting factors for the current mHealth platforms developed for anxiety and depression to become successful solutions. A leap forward would involve venturing into digital therapeutics (DTx), which is widely regarded as the next step for mHealth. DTx offers effective, safe, highly personalized, objective, and cost-saving health care where barriers between information silos are overcome [202], all of which are ideal for addressing mental health conditions such as anxiety and depression. DTx require clinical and regulatory oversight and demonstrated RWE and are further differentiated by the delivery of a clinical-grade intervention (beyond tracking) to treat or manage a disease [198]. In other words, DTx interfere with the patient to modify their state, often through behavioral alteration. They can operate as a standalone treatment (eg, treating a mental disorder through gamification) or adjunct to other therapies to enhance them (eg, improving drug adherence). With the expectation to grow dramatically over the next few years, DTx can potentially affect users, organizations, and the health care industry in different ways: new treatment options, improved care pathways, incorporation of new standards of care, improved patient and population health outcomes, similar coverage to existing therapies, international product quality standards, and acceptance as an independent class of medicine, among others [203]. DTx is a favorable next step in mHealth solutions for anxiety and depression.

Conclusions

This study analyzed the characteristics of mHealth platforms for anxiety or depression that were available for research, commercial use, or both. A systematic search of the current literature and 2 popular app stores identified 169 and 179 mHealth platforms, respectively. mHealth platforms varied in their targeted condition, targeted group, purpose, intervention type, and technology type. In addition, most platforms from popular app stores lacked clinical and RWE, and a small number of platforms present in the literature were available commercially. Future efforts should focus on further accessing the quality—utility, safety, and effectiveness—of the existing platforms and ensuring the adoption by developers, from both commercial and academic sectors alike, of a reporting guideline for their platform description, as well as a regulatory framework to facilitate the development, validation, and deployment of mHealth platforms for anxiety or depression.

Acknowledgments

The authors gratefully acknowledge funding from the Wellcome Trust Mental Health Priority Area Active Ingredients Commission and Institute for Digital Medicine Translational Research Programme (R-719-000-037-733) at the Yong Loo Lin School of

Medicine, the National University of Singapore, and the National Research Foundation Singapore under its AI Singapore Programme (award number: AISG-GC-2019-002) for undertaking this review. The funders had no role in the study design, data collection, management, data interpretation, and writing of this publication.

Conflicts of Interest

DH is a shareholder of KYAN Therapeutics, which has licensed intellectual property pertaining to artificial intelligence–based drug development.

Multimedia Appendix 1
Table of search strings used across the 4 databases and their respective numbers of outcomes.
[DOCX File , 31 KB - jmir_v24i2e27388_app1.docx]

Multimedia Appendix 2
Table of brief mobile health platform characteristics from the literature search.
[DOCX File , 47 KB - jmir_v24i2e27388_app2.docx]

Multimedia Appendix 3
Table of brief mobile health platform characteristics from the app store search.
[DOCX File , 69 KB - jmir_v24i2e27388_app3.docx]

References

1. Frequently asked questions on Global Task Force on digital health for TB and its work. World Health Organization. 2017. URL: https://tinyurl.com/ycknp9en [accessed 2020-12-10]
2. Berrouiguet S, Baca-García E, Brandt S, Walter M, Courtet P. Fundamentals for future mobile-health (mHealth): a systematic review of mobile phone and web-based text messaging in mental health. J Med Internet Res 2016 Dec 10;18(6):e135 [FREE Full text] [doi: 10.2196/jmir.5066] [Medline: 27287668]
3. Walters K. The purpose and value of mHealth for the University of North Carolina at Chapel Hill: a white paper on mHealth. UNC Health Sciences Library. 2014. URL: https://hsl.lib.unc.edu/wp-content/uploads/2016/09/mHealthWhitePaper_1.pdf [accessed 2022-01-10]
4. Rowland SP, Fitzgerald JE, Holme T, Powell J, McGregor A. What is the clinical value of mHealth for patients? NPJ Digit Med 2020 Jan 13;3(1):4 [FREE Full text] [doi: 10.1038/s41746-019-0206-x] [Medline: 31970289]
5. Bradway M, Carrion C, Vallespin B, Saadatfard O, Puigdomènech E, Espallargues M, et al. mHealth assessment: conceptualization of a global framework. JMIR Mhealth Uhealth 2017 May 02;5(5):e60 [FREE Full text] [doi: 10.2196/mhealth.7291] [Medline: 28465282]
6. Price M, Yuen EK, Goetter EM, Herbert JD, Forman EM, Acierno R, et al. mHealth: a mechanism to deliver more accessible, more effective mental health care. Clin Psychol Psychother 2014 Aug;21(5):427-436 [FREE Full text] [doi: 10.1002/cpp.1855] [Medline: 23918764]
7. Kemp S. Digital 2020: October global statshot. DataReportal. 2020. URL: https://datareportal.com/reports/digital-2020-october-global-statshot [accessed 2020-12-10]
8. Digital 2020 in Singapore. We Are Social. 2020. URL: https://wearesocial.com/sg/digital-2020-singapore [accessed 2020-12-10]
9. Number of smartphone users worldwide from 2016 to 2021 (in billions). Newzoo © Statista 2020. 2020. URL: https://tinyurl.com/yjbf8ehw [accessed 2020-12-15]
10. Lecomte T, Potvin S, Corbière M, Guay S, Samson C, Cloutier B, et al. Mobile apps for mental health issues: meta-review of meta-analyses. JMIR Mhealth Uhealth 2020 May 29;8(5):e17458 [FREE Full text] [doi: 10.2196/17458] [Medline: 32348289]
11. Miralles I, Granell C, Díaz-Sanahuja L, Van Woensel W, Bretón-López J, Mira A, et al. Smartphone apps for the treatment of mental disorders: systematic review. JMIR Mhealth Uhealth 2020 Apr 02;8(4):e14897 [FREE Full text] [doi: 10.2196/14897] [Medline: 32238332]
12. Andlin-Sobocki P, Wittchen H. Cost of anxiety disorders in Europe. Eur J Neurol 2005 Jun;12 Suppl 1(s1):39-44. [doi: 10.1111/j.1468-1331.2005.01196.x] [Medline: 15877777]
13. Cuijpers P, Beekman AT, Reynolds CF. Preventing depression: a global priority. J Am Med Assoc 2012 Mar 14;307(10):1033-1034 [FREE Full text] [doi: 10.1001/jama.2012.271] [Medline: 22416097]
14. Sun J, Guo Y, Wang X, Zeng Q. mHealth for aging China: opportunities and challenges. Aging Dis 2016;7(1):53. [doi: 10.14336/ad.2015.1011]
15. Cortez NG, Cohen IG, Kesselheim AS. FDA regulation of mobile health technologies. N Engl J Med 2014 Jul 24;371(4):372-379. [doi: 10.1056/NEJMhle1403384] [Medline: 25054722]

16. Muñoz AO, Camacho E, Torous J. Marketplace and literature review of Spanish language mental health apps. Front Digit Health 2021 Feb 15;3:615366 [FREE Full text] [doi: 10.3389/fdgth.2021.615366] [Medline: 34713093]
17. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
18. Ho GJ, Liew SM, Ng CJ, Shunmugam RH, Glasziou P. Development of a search strategy for an evidence based retrieval service. PLoS One 2016 Dec 9;11(12):e0167170 [FREE Full text] [doi: 10.1371/journal.pone.0167170] [Medline: 27935993]
19. Covidence Systematic Review Software. URL: https://www.covidence.org/ [accessed 2022-01-10]
20. Bender JL, Yue RYK, To MJ, Deacken L, Jadad AR. A lot of action, but not in the right direction: systematic review and content analysis of smartphone applications for the prevention, detection, and management of cancer. J Med Internet Res 2013 Dec;15(12):e287 [FREE Full text] [doi: 10.2196/jmir.2661] [Medline: 24366061]
21. Giunti G, Giunta DH, Guisado-Fernandez E, Bender JL, Fernandez-Luque L. A biopsy of breast cancer mobile applications: state of the practice review. Int J Med Inform 2018 Dec;110:1-9 [FREE Full text] [doi: 10.1016/j.ijmedinf.2017.10.022] [Medline: 29331247]
22. Yokomitsu K, Irie T, Sekiguchi M, Shimizu A, Matsuoka H, Merry SN, et al. Gamified mobile computerized cognitive behavioral therapy for Japanese university students with depressive symptoms: protocol for a randomized controlled trial. JMIR Res Protoc 2020 Apr 07;9(4):e15164 [FREE Full text] [doi: 10.2196/15164] [Medline: 32254045]
23. Reins JA, Ebert DD, Lehr D, Riper H, Cuijpers P, Berking M. Internet-based treatment of major depression for patients on a waiting list for inpatient psychotherapy: protocol for a multi-centre randomised controlled trial. BMC Psychiatry 2013 Nov 26;13(1):318 [FREE Full text] [doi: 10.1186/1471-244X-13-318] [Medline: 24279841]
24. Hirsch A, Luellen J, Holder JM, Steinberg G, Dubiel T, Blazejowskyj A, et al. Managing depressive symptoms in the workplace using a web-based self-care tool: a pilot randomized controlled trial. JMIR Res Protoc 2017 Apr 04;6(4):e51 [FREE Full text] [doi: 10.2196/resprot.7203] [Medline: 28377368]
25. Mira A, Bretón-López J, García-Palacios A, Quero S, Baños RM, Botella C. An Internet-based program for depressive symptoms using human and automated support: a randomized controlled trial. Neuropsychiatr Dis Treat 2017 Mar;Volume 13:987-1006. [doi: 10.2147/ndt.s130994]
26. Ranney ML, Freeman JR, Connell G, Spirito A, Boyer E, Walton M, et al. A depression prevention intervention for adolescents in the Emergency Department. J Adolesc Health 2016 Oct;59(4):401-410 [FREE Full text] [doi: 10.1016/j.jadohealth.2016.04.008] [Medline: 27267141]
27. Robertson L, Smith M, Castle D, Tannenbaum D. Using the internet to enhance the treatment of depression. Australas Psychiat 2006 Dec;14(4):413-417. [doi: 10.1111/j.1440-1665.2006.02315.x]
28. Parker BL, Achilles MR, Subotic-Kerry M, O'Dea B. Youth StepCare: a pilot study of an online screening and recommendations service for depression and anxiety among youth patients in general practice. BMC Fam Pract 2020 Jan 07;21(1):2 [FREE Full text] [doi: 10.1186/s12875-019-1071-z] [Medline: 31910811]
29. Landback J, Prochaska M, Ellis J, Dmochowska K, Kuwabara SA, Gladstone T, et al. From prototype to product: development of a primary care/internet based depression prevention intervention for adolescents (CATCH-IT). Community Ment Health J 2009 Oct 30;45(5):349-354. [doi: 10.1007/s10597-009-9226-3] [Medline: 19641992]
30. Bakker D, Rickard N. Engagement with a cognitive behavioural therapy mobile phone app predicts changes in mental health and wellbeing: MoodMission. Aus Psychol 2020 Nov 12;54(4):245-260. [doi: 10.1111/ap.12383]
31. Hung S, Li M, Chen Y, Chiang J, Chen Y, Hung GC. Smartphone-based ecological momentary assessment for Chinese patients with depression: an exploratory study in Taiwan. Asian J Psychiatr 2016 Oct;23:131-136. [doi: 10.1016/j.ajp.2016.08.003] [Medline: 27969071]
32. Addington EL, Cheung EO, Bassett SM, Kwok I, Schuette SA, Shiu E, et al. The MARIGOLD study: feasibility and enhancement of an online intervention to improve emotion regulation in people with elevated depressive symptoms. J Affect Disord 2019 Oct 01;257:352-364 [FREE Full text] [doi: 10.1016/j.jad.2019.07.049] [Medline: 31302525]
33. Felder J. Mindful mood balance: a case report of web-based treatment of residual depressive symptoms. Perm J 2014 Nov 3:58-62. [doi: 10.7812/tpp/14-019]
34. Garrido S, Cheers D, Boydell K, Nguyen QV, Schubert E, Dunne L, et al. Young people's response to six smartphone apps for anxiety and depression: focus group study. JMIR Ment Health 2019 Oct 02;6(10):e14385 [FREE Full text] [doi: 10.2196/14385] [Medline: 31579023]
35. Szigethy E, Solano F, Wallace M, Perry DL, Morrell L, Scott K, et al. A study protocol for a non-randomised comparison trial evaluating the feasibility and effectiveness of a mobile cognitive-behavioural programme with integrated coaching for anxious adults in primary care. BMJ Open 2018 Dec 13;8(1):e019108 [FREE Full text] [doi: 10.1136/bmjopen-2017-019108] [Medline: 29331971]
36. Witlox M, Kraaij V, Garnefski N, de Waal MW, Smit F, Hoencamp E, et al. An internet-based Acceptance and Commitment Therapy intervention for older adults with anxiety complaints: study protocol for a cluster randomized controlled trial. Trials 2018 Sep 17;19(1):502 [FREE Full text] [doi: 10.1186/s13063-018-2731-3] [Medline: 30223873]

37. van Ballegooijen W, Ruwaard J, Karyotaki E, Ebert DD, Smit JH, Riper H. Reactivity to smartphone-based ecological momentary assessment of depressive symptoms (MoodMonitor): protocol of a randomised controlled trial. BMC Psychiatry 2016 Oct 21;16(1):359 [FREE Full text] [doi: 10.1186/s12888-016-1065-5] [Medline: 27769201]
38. Stiles-Shields C, Montague E, Kwasny MJ, Mohr DC. Behavioral and cognitive intervention strategies delivered via coached apps for depression: pilot trial. Psychol Serv 2019 May;16(2):233-238 [FREE Full text] [doi: 10.1037/ser0000261] [Medline: 30407055]
39. Fleming JB, Hill YN, Burns MN. Usability of a culturally informed mHealth intervention for symptoms of anxiety and depression: feedback from young sexual minority men. JMIR Hum Factors 2017 Aug 25;4(3):e22 [FREE Full text] [doi: 10.2196/humanfactors.7392] [Medline: 28842389]
40. Chung K, Park JY, Joung D, Jhung K. Response time as an implicit self-schema indicator for depression among undergraduate students: preliminary findings from a mobile app-based depression assessment. JMIR Mhealth Uhealth 2019 Sep 13;7(9):e14657 [FREE Full text] [doi: 10.2196/14657] [Medline: 31586362]
41. Stoll RD, Pina AA, Gary K, Amresh A. Usability of a smartphone application to support the prevention and early intervention of anxiety in youth. Cogn Behav Pract 2017 Nov;24(4):393-404 [FREE Full text] [doi: 10.1016/j.cbpra.2016.11.002] [Medline: 29056845]
42. Meglic M, Furlan M, Kuzmanic M, Kozel D, Baraga D, Kuhar I, et al. Feasibility of an eHealth service to support collaborative depression care: results of a pilot study. J Med Internet Res 2010 Dec 19;12(5):e63 [FREE Full text] [doi: 10.2196/jmir.1510] [Medline: 21172765]
43. Nakao S, Nakagawa A, Oguchi Y, Mitsuda D, Kato N, Nakagawa Y, et al. Web-based cognitive behavioral therapy blended with face-to-face sessions for major depression: randomized controlled trial. J Med Internet Res 2018 Dec 21;20(9):e10743 [FREE Full text] [doi: 10.2196/10743] [Medline: 30249583]
44. Moeini B, Bashirian S, Soltanian A, Ghaleiha A, Taheri M. Examining the effectiveness of a web-based intervention for depressive symptoms in female adolescents: applying social cognitive theory. J Res Health Sci 2019;19(3). [doi: 10.34172/jrhs194826]
45. Berger T, Krieger T, Sude K, Meyer B, Maercker A. Evaluating an e-mental health program ("deprexis") as adjunctive treatment tool in psychotherapy for depression: results of a pragmatic randomized controlled trial. J Affect Disord 2018 Dec;227:455-462. [doi: 10.1016/j.jad.2017.11.021] [Medline: 29154168]
46. Pinto MD, Greenblatt AM, Hickman RL, Rice HM, Thomas TL, Clochesy JM. Assessing the critical parameters of eSMART-MH: a promising avatar-based digital therapeutic intervention to reduce depressive symptoms. Perspect Psychiatr Care 2016 Jul 19;52(3):157-168. [doi: 10.1111/ppc.12112] [Medline: 25800698]
47. MacLean S, Corsi DJ, Litchfield S, Kucharski J, Genise K, Selaman Z, et al. Coach-facilitated web-based therapy compared with information about web-based resources in patients referred to secondary mental health care for depression: randomized controlled trial. J Med Internet Res 2020 Jun 09;22(6):e15001 [FREE Full text] [doi: 10.2196/15001] [Medline: 32515740]
48. Geraedts AS, Kleiboer AM, Wiezer NM, Cuijpers P, van Mechelen W, Anema JR. Feasibility of a worker-directed web-based intervention for employees with depressive symptoms. Internet Interven 2014 Jul;1(3):132-140. [doi: 10.1016/j.invent.2014.07.001]
49. Christensen H, Griffiths KM, Korten A. Web-based cognitive behavior therapy: analysis of site usage and changes in depression and anxiety scores. J Med Internet Res 2002 Feb;4(1):e3 [FREE Full text] [doi: 10.2196/jmir.4.1.e3] [Medline: 11956035]
50. Goldin PR, Lindholm R, Ranta K, Hilgert O, Helteenvuori T, Raevuori A. Feasibility of a therapist-supported, mobile phone-delivered online intervention for depression: longitudinal observational study. JMIR Form Res 2019 Jan 22;3(1):e11509 [FREE Full text] [doi: 10.2196/11509] [Medline: 30682726]
51. Karyotaki E, Klein AM, Riper H, Wit LD, Krijnen L, Bol E, et al. Examining the effectiveness of a web-based intervention for symptoms of depression and anxiety in college students: study protocol of a randomised controlled trial. BMJ Open 2019;9(5):e028739 [FREE Full text] [doi: 10.1136/bmjopen-2018-028739] [Medline: 31092668]
52. Geraedts AS, Kleiboer AM, Wiezer NM, van Mechelen W, Cuijpers P. Web-based guided self-help for employees with depressive symptoms (Happy@Work): design of a randomized controlled trial. BMC Psychiatry 2013 Feb 18;13(1). [doi: 10.1186/1471-244x-13-61]
53. Lopes RT, Meyer B, Berger T, Svacina MA. Effectiveness of an internet-based self-guided program to treat depression in a sample of Brazilian users: a study protocol. Braz J Psychiatry 2020 Jun;42(3):322-328 [FREE Full text] [doi: 10.1590/1516-4446-2019-0582] [Medline: 32159711]
54. Pramana G, Parmanto B, Kendall PC, Silk JS. The SmartCAT: an m-health platform for ecological momentary intervention in child anxiety treatment. Telemed J E Health 2014 May;20(5):419-427 [FREE Full text] [doi: 10.1089/tmj.2013.0214] [Medline: 24579913]
55. Yap MB, Mahtani S, Rapee RM, Nicolas C, Lawrence KA, Mackinnon A, et al. A tailored web-based intervention to improve parenting risk and protective factors for adolescent depression and anxiety problems: postintervention findings from a randomized controlled trial. J Med Internet Res 2018 Dec 19;20(1):e17 [FREE Full text] [doi: 10.2196/jmir.9139] [Medline: 29351895]

56. Fulgoni CM, Melvin GA, Jorm AF, Lawrence KA, Yap MB. The Therapist-assisted Online Parenting Strategies (TOPS) program for parents of adolescents with clinical anxiety or depression: development and feasibility pilot. Internet Interv 2019 Dec;18:100285 [FREE Full text] [doi: 10.1016/j.invent.2019.100285] [Medline: 31890632]
57. Kladnitski N, Smith J, Uppal S, James MA, Allen AR, Andrews G, et al. Transdiagnostic internet-delivered CBT and mindfulness-based treatment for depression and anxiety: a randomised controlled trial. Internet Interv 2020 Apr;20:100310 [FREE Full text] [doi: 10.1016/j.invent.2020.100310] [Medline: 32140427]
58. Farrer L, Christensen H, Griffiths KM, Mackinnon A. Internet-based CBT for depression with and without telephone tracking in a national helpline: randomised controlled trial. PLoS One 2011 Nov;6(11):e28099 [FREE Full text] [doi: 10.1371/journal.pone.0028099] [Medline: 22140514]
59. Jolstedt M, Ljótsson B, Fredlander S, Tedgård T, Hallberg A, Ekeljung A, et al. Implementation of internet-delivered CBT for children with anxiety disorders in a rural area: a feasibility trial. Internet Interv 2018 Jun;12:121-129 [FREE Full text] [doi: 10.1016/j.invent.2017.11.003] [Medline: 30135776]
60. Berger T, Hämmerli K, Gubser N, Andersson G, Caspar F. Internet-based treatment of depression: a randomized controlled trial comparing guided with unguided self-help. Cogn Behav Ther 2011 Dec;40(4):251-266. [doi: 10.1080/16506073.2011.616531] [Medline: 22060248]
61. Moberg C, Niles A, Beermann D. Guided self-help works: randomized waitlist controlled trial of Pacifica, a mobile app integrating cognitive behavioral therapy and mindfulness for stress, anxiety, and depression. J Med Internet Res 2019 Jun 08;21(6):e12556 [FREE Full text] [doi: 10.2196/12556] [Medline: 31199319]
62. Dimidjian S, Beck A, Felder JN, Boggs JM, Gallop R, Segal ZV. Web-based Mindfulness-based Cognitive Therapy for reducing residual depressive symptoms: an open trial and quasi-experimental comparison to propensity score matched controls. Behav Res Ther 2014 Dec;63:83-89 [FREE Full text] [doi: 10.1016/j.brat.2014.09.004] [Medline: 25461782]
63. Moukaddam N, Truong A, Cao J, Shah A, Sabharwal A. Findings from a trial of the smartphone and online usage-based evaluation for depression (SOLVD) application: what do apps really tell us about patients with depression? Concordance between app-generated data and standard psychiatric questionnaires for depression and anxiety. J Psychiatr Pract 2019 Sep;25(5):365-373. [doi: 10.1097/PRA.0000000000000420] [Medline: 31505521]
64. Kamath J, Bi J, Russell A, Wang B. Grant report on SCH: personalized depression treatment supported by mobile sensor analytics. J Psychiatr Brain Sci 2020;5 [FREE Full text] [doi: 10.20900/jpbs.20200010] [Medline: 32529036]
65. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. JMIR Ment Health 2017 Jun 06;4(2):e19 [FREE Full text] [doi: 10.2196/mental.7785] [Medline: 28588005]
66. Geraedts AS, Kleiboer AM, Twisk J, Wiezer NM, van Mechelen W, Cuijpers P. Long-term results of a web-based guided self-help intervention for employees with depressive symptoms: randomized controlled trial. J Med Internet Res 2014 Jul;16(7):e168 [FREE Full text] [doi: 10.2196/jmir.3539] [Medline: 25008127]
67. El Morr C, Ritvo P, Ahmad F, Moineddin R, Team M. Correction: effectiveness of an 8-week web-based mindfulness virtual community intervention for university students on symptoms of stress, anxiety, and depression: randomized controlled trial. JMIR Ment Health 2020 Sep 30;7(9):e24131 [FREE Full text] [doi: 10.2196/24131] [Medline: 32997643]
68. Watts S, Mackenzie A, Thomas C, Griskaitis A, Mewton L, Williams A, et al. CBT for depression: a pilot RCT comparing mobile phone vs. computer. BMC Psychiatry 2013;13:49 [FREE Full text] [doi: 10.1186/1471-244X-13-49] [Medline: 23391304]
69. Watanabe N, Horikoshi M, Yamada M, Shimodera S, Akechi T, Miki K, Steering Committee of the Fun to Learn to ActThink through Technology Project. Adding smartphone-based cognitive-behavior therapy to pharmacotherapy for major depression (FLATT project): study protocol for a randomized controlled trial. Trials 2015 Jul 07;16:293 [FREE Full text] [doi: 10.1186/s13063-015-0805-z] [Medline: 26149441]
70. Silk JS, Pramana G, Sequeira SL, Lindhiem O, Kendall PC, Rosen D, et al. Using a smartphone app and clinician portal to enhance brief cognitive behavioral therapy for childhood anxiety disorders. Behav Ther 2020 Jan;51(1):69-84 [FREE Full text] [doi: 10.1016/j.beth.2019.05.002] [Medline: 32005341]
71. Farvolden P, McBride C, Bagby RM, Ravitz P. A Web-based screening instrument for depression and anxiety disorders in primary care. J Med Internet Res 2003 Sep 29;5(3):e23 [FREE Full text] [doi: 10.2196/jmir.5.3.e23] [Medline: 14517114]
72. Massoudi B, Blanker MH, van Valen E, Wouters H, Bockting CL, Burger H. Blended care vs. usual care in the treatment of depressive symptoms and disorders in general practice [BLENDING]: study protocol of a non-inferiority randomized trial. BMC Psychiatry 2017 Dec 13;17(1):218 [FREE Full text] [doi: 10.1186/s12888-017-1376-1] [Medline: 28610561]
73. Lara MA, Tiburcio M, Abrego AA, Sánchez-Solís A. A four-year experience with a web-based self-help intervention for depressive symptoms in Mexico. Rev Panam Salud Public 2014;35(5-6):399-406. [Medline: 25211568]
74. Fortney J, Pyne J, Steven C, Williams JS, Hedrick RG, Lunsford AK, et al. A web-based clinical decision support system for depression care management. Am J Manag Care 2010 Nov;16(11):849-854 [FREE Full text] [Medline: 21348556]
75. Ünlü Ince B, Gökçay D, Riper H, Cuijpers P. Effectiveness of a web- and mobile-guided psychological intervention for depressive symptoms in turkey: protocol for a randomized controlled trial. JMIR Res Protoc 2019 Apr 05;8(4):e13239 [FREE Full text] [doi: 10.2196/13239] [Medline: 30950802]

76. Høifødt RS, Mittner M, Lillevoll K, Katla SK, Kolstrup N, Eisemann M, et al. Predictors of response to web-based cognitive behavioral therapy with high-intensity face-to-face therapist guidance for depression: a Bayesian analysis. J Med Internet Res 2015 Sep 02;17(9):e197 [FREE Full text] [doi: 10.2196/jmir.4351] [Medline: 26333818]
77. Pratap A, Renn BN, Volponi J, Mooney SD, Gazzaley A, Arean PA, et al. Using mobile apps to assess and treat depression in Hispanic and Latino populations: fully remote randomized clinical trial. J Med Internet Res 2018 Aug 09;20(8):e10130 [FREE Full text] [doi: 10.2196/10130] [Medline: 30093372]
78. Cheung EO, Addington EL, Bassett SM, Schuette SA, Shiu EW, Cohn MA, et al. A self-paced, web-based, positive emotion skills intervention for reducing symptoms of depression: protocol for development and pilot testing of MARIGOLD. JMIR Res Protoc 2018 Jun 05;7(6):e10494 [FREE Full text] [doi: 10.2196/10494] [Medline: 29871853]
79. Schneider BC, Schröder J, Berger T, Hohagen F, Meyer B, Späth C, et al. Bridging the "digital divide": a comparison of use and effectiveness of an online intervention for depression between Baby Boomers and Millennials. J Affect Disord 2018 Aug 15;236:243-251. [doi: 10.1016/j.jad.2018.04.101] [Medline: 29751239]
80. Buntrock C, Ebert DD, Lehr D, Cuijpers P, Riper H, Smit F, et al. Evaluating the efficacy and cost-effectiveness of web-based indicated prevention of major depression: design of a randomised controlled trial. BMC Psychiatry 2014 Jan 31;14(1). [doi: 10.1186/1471-244x-14-25]
81. Dahne J, Lejuez CW, Kustanowitz J, Felton JW, Diaz VA, Player MS, et al. Moodivate: a self-help behavioral activation mobile app for utilization in primary care - Development and clinical considerations. Int J Psychiatry Med 2017 Mar;52(2):160-175 [FREE Full text] [doi: 10.1177/0091217417720899] [Medline: 28792292]
82. Yap MB, Cardamone-Breen MC, Rapee RM, Lawrence KA, Mackinnon AJ, Mahtani S, et al. Medium-term effects of a tailored web-based parenting intervention to reduce adolescent risk of depression and anxiety: 12-month findings from a randomized controlled trial. J Med Internet Res 2019 Aug 15;21(8):e13628. [doi: 10.2196/13628]
83. Rice S, O'Bree B, Wilson M, McEnery C, Lim MH, Hamilton M, et al. Leveraging the social network for treatment of social anxiety: pilot study of a youth-specific digital intervention with a focus on engagement of young men. Internet Interv 2020 Apr;20:100323 [FREE Full text] [doi: 10.1016/j.invent.2020.100323] [Medline: 32435600]
84. Economides M, Ranta K, Nazander A, Hilgert O, Goldin PR, Raevuori A, et al. Long-term outcomes of a therapist-supported, smartphone-based intervention for elevated symptoms of depression and anxiety: quasiexperimental, pre-postintervention study. JMIR Mhealth Uhealth 2019 Aug 26;7(8):e14284 [FREE Full text] [doi: 10.2196/14284] [Medline: 31452521]
85. Bücker L, Schnakenberg P, Karyotaki E, Moritz S, Westermann S. Diminishing effects after recurrent use of self-guided internet-based interventions in depression: randomized controlled trial. J Med Internet Res 2019 Oct 02;21(10):e14240 [FREE Full text] [doi: 10.2196/14240] [Medline: 31579014]
86. Ranney ML, Patena JV, Dunsiger S, Spirito A, Cunningham RM, Boyer E, et al. A technology-augmented intervention to prevent peer violence and depressive symptoms among at-risk emergency department adolescents: protocol for a randomized control trial. Contemp Clin Trials 2019 Jul;82:106-114 [FREE Full text] [doi: 10.1016/j.cct.2019.05.009] [Medline: 31129373]
87. Whitton AE, Proudfoot J, Clarke J, Birch M, Parker G, Manicavasagar V, et al. Breaking open the black box: isolating the most potent features of a web and mobile phone-based intervention for depression, anxiety, and stress. JMIR Ment Health 2015 Mar 04;2(1):e3 [FREE Full text] [doi: 10.2196/mental.3573] [Medline: 26543909]
88. Christensen H, Mackinnon AJ, Batterham PJ, O'Dea B, Guastella AJ, Griffiths KM, et al. The effectiveness of an online e-health application compared to attention placebo or Sertraline in the treatment of Generalised Anxiety Disorder. Internet Interv 2014 Oct;1(4):169-174. [doi: 10.1016/j.invent.2014.08.002]
89. Pramana G, Parmanto B, Lomas J, Lindhiem O, Kendall PC, Silk J. Using mobile health gamification to facilitate cognitive behavioral therapy skills practice in child anxiety treatment: open clinical trial. JMIR Serious Games 2018 May 10;6(2):e9 [FREE Full text] [doi: 10.2196/games.8902] [Medline: 29748165]
90. Proudfoot J, Clarke J, Birch M, Whitton AE, Parker G, Manicavasagar V, et al. Impact of a mobile phone and web program on symptom and functional outcomes for people with mild-to-moderate depression, anxiety and stress: a randomised controlled trial. BMC Psychiatry 2013;13:312 [FREE Full text] [doi: 10.1186/1471-244X-13-312] [Medline: 24237617]
91. Bae J, Wolpin S, Kim E, Lee S, Yoon S, An K. Development of a user-centered health information service system for depressive symptom management. Nurs Health Sci 2009 Jun;11(2):185-193. [doi: 10.1111/j.1442-2018.2009.00454.x] [Medline: 19519707]
92. Mohr DC, Tomasino KN, Lattie EG, Palac HL, Kwasny MJ, Weingardt K, et al. IntelliCare: an eclectic, skills-based app suite for the treatment of depression and anxiety. J Med Internet Res 2017 Jan 05;19(1):e10 [FREE Full text] [doi: 10.2196/jmir.6645] [Medline: 28057609]
93. Cardamone-Breen MC, Jorm AF, Lawrence KA, Rapee RM, Mackinnon AJ, Yap MB. A single-session, web-based parenting intervention to prevent adolescent depression and anxiety disorders: randomized controlled trial. J Med Internet Res 2018 Apr 26;20(4):e148 [FREE Full text] [doi: 10.2196/jmir.9499] [Medline: 29699964]
94. Kleiboer A, Donker T, Seekles W, van Straten A, Riper H, Cuijpers P. A randomized controlled trial on the role of support in internet-based problem solving therapy for depression and anxiety. Behav Res Ther 2015 Sep;72:63-71. [doi: 10.1016/j.brat.2015.06.013] [Medline: 26188373]

95. Beiwinkel T, Eißing T, Telle N, Siegmund-Schultze E, Rössler W. Effectiveness of a web-based intervention in reducing depression and sickness absence: randomized controlled trial. J Med Internet Res 2017 Jun 15;19(6):e213 [FREE Full text] [doi: 10.2196/jmir.6546] [Medline: 28619701]
96. Corden ME, Koucky EM, Brenner C, Palac HL, Soren A, Begale M, et al. MedLink: a mobile intervention to improve medication adherence and processes of care for treatment of depression in general medicine. Digit Health 2016 Aug 03;2:2055207616663069 [FREE Full text] [doi: 10.1177/2055207616663069] [Medline: 29942564]
97. Roy-Byrne P, Craske MG, Sullivan G, Rose RD, Edlund MJ, Lang AJ, et al. Delivery of evidence-based treatment for multiple anxiety disorders in primary care: a randomized controlled trial. J Am Med Assoc 2010 May 19;303(19):1921-1928 [FREE Full text] [doi: 10.1001/jama.2010.608] [Medline: 20483968]
98. Mohr D, Stiles-Shields C, Brenner C, Palac H, Montague E, Kaiser SM, et al. MedLink: a mobile intervention to address failure points in the treatment of depression in general medicine. Int Conf Pervasive Comput Technol Healthc 2015 May;2015:100-107 [FREE Full text] [doi: 10.4108/icst.pervasivehealth.2015.259042] [Medline: 26640740]
99. March S, Spence SH, Donovan CL, Kenardy JA. Large-scale dissemination of internet-based cognitive behavioral therapy for youth anxiety: feasibility and acceptability study. J Med Internet Res 2018 Jul 04;20(7):e234 [FREE Full text] [doi: 10.2196/jmir.9211] [Medline: 29973338]
100. Narziev N, Goh H, Toshnazarov K, Lee SA, Chung K, Noh Y. STDD: Short-Term Depression Detection with Passive Sensing. Sensors (Basel) 2020 Mar 04;20(5):1396 [FREE Full text] [doi: 10.3390/s20051396] [Medline: 32143358]
101. Hobbs MJ, Mahoney AE, Andrews G. Integrating iCBT for generalized anxiety disorder into routine clinical care: treatment effects across the adult lifespan. J Anxiety Disord 2017 Oct;51:47-54. [doi: 10.1016/j.janxdis.2017.09.003] [Medline: 28926805]
102. Hatcher S, Whittaker R, Patton M, Miles WS, Ralph N, Kercher K, et al. Web-based therapy plus support by a coach in depressed patients referred to secondary mental health care: randomized controlled trial. JMIR Ment Health 2018 Jan 23;5(1):e5 [FREE Full text] [doi: 10.2196/mental.8510] [Medline: 29362207]
103. Cook L, Mostazir M, Watkins E. Reducing Stress and Preventing Depression (RESPOND): randomized controlled trial of web-based rumination-focused cognitive behavioral therapy for high-ruminating university students. J Med Internet Res 2019 May 13;21(5):e11349 [FREE Full text] [doi: 10.2196/11349] [Medline: 31094331]
104. Meyer B, Berger T, Caspar F, Beevers CG, Andersson G, Weiss M. Effectiveness of a novel integrative online treatment for depression (Deprexis): randomized controlled trial. J Med Internet Res 2009 May 11;11(2):e15 [FREE Full text] [doi: 10.2196/jmir.1151] [Medline: 19632969]
105. Ramos RM, Cheng PG, Jonas SM. Validation of an mHealth app for depression screening and monitoring (Psychologist in a Pocket): correlational study and concurrence analysis. JMIR Mhealth Uhealth 2019 Sep 16;7(9):e12051 [FREE Full text] [doi: 10.2196/12051] [Medline: 31538946]
106. Dang M, Mielke C, Diehl A, Haux R. Accompanying depression with FINE - A smartphone-based approach. In: Studies in Health Technology and Informatics. Amsterdam: IOS Press; 2016:195-199.
107. Kwasny MJ, Schueller SM, Lattie E, Gray EL, Mohr DC. Exploring the use of multiple mental health apps within a platform: secondary analysis of the IntelliCare field trial. JMIR Ment Health 2019 Mar 21;6(3):e11572 [FREE Full text] [doi: 10.2196/11572] [Medline: 30896433]
108. Deady M, Johnston DA, Glozier N, Milne D, Choi I, Mackinnon A, et al. Smartphone application for preventing depression: study protocol for a workplace randomised controlled trial. BMJ Open 2018 Jul 13;8(7):e020510 [FREE Full text] [doi: 10.1136/bmjopen-2017-020510] [Medline: 30007927]
109. Gaudiano BA, Davis CH, Miller IW, Uebelacker L. Pilot randomized controlled trial of a video self-help intervention for depression based on acceptance and commitment therapy: feasibility and acceptability. Clin Psychol Psychother 2020 May 28;27(3):396-407 [FREE Full text] [doi: 10.1002/cpp.2436] [Medline: 32087610]
110. Lintvedt OK, Griffiths KM, Sørensen K, Østvik AR, Wang CE, Eisemann M, et al. Evaluating the effectiveness and efficacy of unguided internet-based self-help intervention for the prevention of depression: a randomized controlled trial. Clin Psychol Psychother 2013 Sep 02;20(1):10-27. [doi: 10.1002/cpp.770] [Medline: 21887811]
111. Kemmeren LL, van Schaik DJ, Riper H, Kleiboer AM, Bosmans JE, Smit JH. Effectiveness of blended depression treatment for adults in specialised mental healthcare: study protocol for a randomised controlled trial. BMC Psychiatry 2016 Apr 21;16:113 [FREE Full text] [doi: 10.1186/s12888-016-0818-5] [Medline: 27102812]
112. Radovic A, Li Y, Landsittel D, Stein BD, Miller E. A social media website (supporting our valued adolescents) to support treatment uptake for adolescents with depression and/or anxiety and their parents: protocol for a pilot randomized controlled trial. JMIR Res Protoc 2019 Jan 23;8(1):e12117 [FREE Full text] [doi: 10.2196/12117] [Medline: 30672741]
113. Kelders SM, Bohlmeijer ET, Van Gemert-Pijnen JE. Participants, usage, and use patterns of a web-based intervention for the prevention of depression within a randomized controlled trial. J Med Internet Res 2013 Aug;15(8):e172 [FREE Full text] [doi: 10.2196/jmir.2258] [Medline: 23963284]
114. Kramer J, Conijn B, Oijevaar P, Riper H. Effectiveness of a web-based solution-focused brief chat treatment for depressed adolescents and young adults: randomized controlled trial. J Med Internet Res 2014;16(5):e141 [FREE Full text] [doi: 10.2196/jmir.3261] [Medline: 24874006]

115. Buntrock C, Ebert D, Lehr D, Riper H, Smit F, Cuijpers P, et al. Effectiveness of a web-based cognitive behavioural intervention for subthreshold depression: pragmatic randomised controlled trial. Psychother Psychosom 2015 Sep;84(6):348-358. [doi: 10.1159/000438673] [Medline: 26398885]
116. Zwerenz R, Baumgarten C, Becker J, Tibubos A, Siepmann M, Knickenberg RJ, et al. Improving the course of depressive symptoms after inpatient psychotherapy using adjunct web-based self-help: follow-up results of a randomized controlled trial. J Med Internet Res 2019 Oct 24;21(10):e13655 [FREE Full text] [doi: 10.2196/13655] [Medline: 31651403]
117. Clarke G, Reid E, Eubanks D, O'Connor E, DeBar LL, Kelleher C, et al. Overcoming depression on the internet (ODIN): a randomized controlled trial of an Internet depression skills intervention program. J Med Internet Res 2002 Dec;4(3):E14 [FREE Full text] [doi: 10.2196/jmir.4.3.e14] [Medline: 12554545]
118. Ponzo S, Morelli D, Kawadler JM, Hemmings NR, Bird G, Plans D. Efficacy of the digital therapeutic mobile app biobase to reduce stress and improve mental well-being among university students: randomized controlled trial. JMIR Mhealth Uhealth 2020 Apr 06;8(4):e17767 [FREE Full text] [doi: 10.2196/17767] [Medline: 31926063]
119. Arean PA, Hallgren KA, Jordan JT, Gazzaley A, Atkins DC, Heagerty PJ, et al. The use and effectiveness of mobile apps for depression: results from a fully remote clinical trial. J Med Internet Res 2016 Dec 20;18(12):e330 [FREE Full text] [doi: 10.2196/jmir.6482] [Medline: 27998876]
120. Unlü B, Riper H, van SA, Cuijpers P. Guided self-help on the internet for Turkish migrants with depression: the design of a randomized controlled trial. Trials 2010;11:101 [FREE Full text] [doi: 10.1186/1745-6215-11-101] [Medline: 21047442]
121. Deady M, Johnston D, Milne D, Glozier N, Peters D, Calvo R, et al. Preliminary effectiveness of a smartphone app to reduce depressive symptoms in the workplace: feasibility and acceptability study. JMIR Mhealth Uhealth 2018 Dec 04;6(12):e11661 [FREE Full text] [doi: 10.2196/11661] [Medline: 30514694]
122. Palacios JE, Richards D, Palmer R, Coudray C, Hofmann SG, Palmieri PA, et al. Supported internet-delivered cognitive behavioral therapy programs for depression, anxiety, and stress in university students: open, non-randomised trial of acceptability, effectiveness, and satisfaction. JMIR Ment Health 2018 Dec 14;5(4):e11467 [FREE Full text] [doi: 10.2196/11467] [Medline: 30552076]
123. Duraimani S. A Cross-sectional and longitudinal study of the effects of a mindfulness meditation mobile application platform on reducing stress and anxiety. Int J Yoga 2019;12(3):226. [doi: 10.4103/ijoy.ijoy_56_18]
124. Bakker D, Kazantzis N, Rickwood D, Rickard N. Development and pilot evaluation of smartphone-delivered cognitive behavior therapy strategies for mood- and anxiety-related problems: MoodMission. Cogn Behav Pract 2018 Nov;25(4):496-514. [doi: 10.1016/j.cbpra.2018.07.002]
125. Trivedi M, Jha M, Kahalnik F, Pipes R, Levinson S, Lawson T, et al. VitalSign: a Primary Care First (PCP-First) Model for universal screening and measurement-based care for depression. Pharmaceuticals (Basel) 2019 May 14;12(2):71 [FREE Full text] [doi: 10.3390/ph12020071] [Medline: 31091770]
126. Bevan Jones R, Thapar A, Rice F, Mars B, Agha SS, Smith D, et al. A digital intervention for adolescent depression (MoodHwb): mixed methods feasibility evaluation. JMIR Ment Health 2020 Jul 17;7(7):e14536 [FREE Full text] [doi: 10.2196/14536] [Medline: 32384053]
127. Lappalainen P, Langrial S, Oinas-Kukkonen H, Tolvanen A, Lappalainen R. Web-based acceptance and commitment therapy for depressive symptoms with minimal support: a randomized controlled trial. Behav Modif 2015 Nov;39(6):805-834. [doi: 10.1177/0145445515598142] [Medline: 26253644]
128. Pham Q, Khatib Y, Stansfeld S, Fox S, Green T. Feasibility and efficacy of an mhealth game for managing anxiety: "Flowy" randomized controlled pilot trial and design evaluation. Games Health J 2016 Feb;5(1):50-67. [doi: 10.1089/g4h.2015.0033] [Medline: 26536488]
129. Mohr DC, Duffecy J, Ho J, Kwasny M, Cai X, Burns MN, et al. A randomized controlled trial evaluating a manualized TeleCoaching protocol for improving adherence to a web-based intervention for the treatment of depression. PLoS One 2013 Aug;8(8):e70086 [FREE Full text] [doi: 10.1371/journal.pone.0070086] [Medline: 23990896]
130. Lüdtke T, Pult LK, Schröder J, Moritz S, Bücker L. A randomized controlled trial on a smartphone self-help application (Be Good to Yourself) to reduce depressive symptoms. Psychiatry Res 2018 Nov;269:753-762. [doi: 10.1016/j.psychres.2018.08.113] [Medline: 30273901]
131. Clarke J, Proudfoot J, Whitton A, Birch M, Boyd M, Parker G, et al. Therapeutic alliance with a fully automated mobile phone and web-based intervention: secondary analysis of a randomized controlled trial. JMIR Ment Health 2016 Feb 25;3(1):e10 [FREE Full text] [doi: 10.2196/mental.4656] [Medline: 26917096]
132. Chen RY, Feltes JR, Tzeng WS, Lu ZY, Pan M, Zhao N, et al. Phone-based interventions in adolescent psychiatry: a perspective and proof of concept pilot study with a focus on depression and autism. JMIR Res Protoc 2017 Jun 16;6(6):e114 [FREE Full text] [doi: 10.2196/resprot.7245] [Medline: 28623183]
133. Lambert JD, Greaves CJ, Farrand P, Price L, Haase AM, Taylor AH. Web-based intervention using behavioral activation and physical activity for adults with depression (The eMotion Study): pilot randomized controlled trial. J Med Internet Res 2018 Jul 16;20(7):e10112 [FREE Full text] [doi: 10.2196/10112] [Medline: 30012547]
134. Howell AN, Rheingold AA, Uhde TW, Guille C. Web-based CBT for the prevention of anxiety symptoms among medical and health science graduate students. Cogn Behav Ther 2019 Sep 11;48(5):385-405 [FREE Full text] [doi: 10.1080/16506073.2018.1533575] [Medline: 30526364]

135. Wisniewski SR, Eng H, Meloro L, Gatt R, Ritz L, Stegman D, et al. Web-based communications and management of a multi-center clinical trial: the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) project. Clin Trials 2004 Sep 03;1(4):387-398. [doi: 10.1191/1740774504cn035oa] [Medline: 16279277]
136. Fulmer R, Joerin A, Gentile B, Lakerink L, Rauws M. Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: randomized controlled trial. JMIR Ment Health 2018 Dec 13;5(4):e64 [FREE Full text] [doi: 10.2196/mental.9782] [Medline: 30545815]
137. Wahle F, Kowatsch T, Fleisch E, Rufer M, Weidt S. Mobile sensing and support for people with depression: a pilot trial in the wild. JMIR Mhealth Uhealth 2016 Sep 21;4(3):e111 [FREE Full text] [doi: 10.2196/mhealth.5960] [Medline: 27655245]
138. Burton C, Tatar A, McKinstry B, Matheson C, Matu S, Moldovan R, et al. Pilot randomised controlled trial of Help4Mood, an embodied virtual agent-based system to support treatment of depression. J Telemed Telecare 2016 Jul 09;22(6):348-355. [doi: 10.1177/1357633x15609793]
139. Kordy H, Wolf M, Aulich K, Bürgy M, Hegerl U, Hüsing J, et al. Internet-delivered disease management for recurrent depression: a multicenter randomized controlled trial. Psychother Psychosom 2016 Jan;85(2):91-98. [doi: 10.1159/000441951] [Medline: 26808817]
140. Bitsch J, Ramos R, Ix T, Ferrer-Cheng P, Wehrle K. Psychologist in a pocket: towards depression screening on mobile phones. In: Studies in Health Technology and Informatics. Amsterdam: IOS Press; 2015:153-159.
141. Torous J, Staples P, Shanahan M, Lin C, Peck P, Keshavan M, et al. Utilizing a personal smartphone custom app to assess the patient health questionnaire-9 (PHQ-9) depressive symptoms in patients with major depressive disorder. JMIR Ment Health 2015 Mar;2(1):e8 [FREE Full text] [doi: 10.2196/mental.3889] [Medline: 26543914]
142. Mohr DC, Schueller SM, Tomasino KN, Kaiser SM, Alam N, Karr C, et al. Comparison of the effects of coaching and receipt of app recommendations on depression, anxiety, and engagement in the IntelliCare platform: factorial randomized controlled trial. J Med Internet Res 2019 Aug 28;21(8):e13609 [FREE Full text] [doi: 10.2196/13609] [Medline: 31464192]
143. Jones RB, Thapar A, Rice F, Beeching H, Cichosz R, Mars B, et al. A web-based psychoeducational intervention for adolescent depression: design and development of MoodHwb. JMIR Ment Health 2018 Feb 15;5(1):e13 [FREE Full text] [doi: 10.2196/mental.8894] [Medline: 29449202]
144. Furukawa TA, Horikoshi M, Fujita H, Tsujino N, Jinnin R, Kako Y, et al. Cognitive and behavioral skills exercises completed by patients with major depression during smartphone cognitive behavioral therapy: secondary analysis of a randomized controlled trial. JMIR Ment Health 2018 Jan 11;5(1):e4 [FREE Full text] [doi: 10.2196/mental.9092] [Medline: 29326098]
145. Aguilera A, Bruehlman-Senecal E, Demasi O, Avila P. Automated text messaging as an adjunct to cognitive behavioral therapy for depression: a clinical trial. J Med Internet Res 2017 May 08;19(5):e148 [FREE Full text] [doi: 10.2196/jmir.6914] [Medline: 28483742]
146. Christensen H, Guastella AJ, Mackinnon AJ, Griffiths KM, Eagleson C, Batterham PJ, et al. Protocol for a randomised controlled trial investigating the effectiveness of an online e-health application compared to attention placebo or sertraline in the treatment of generalised anxiety disorder. Trials 2010 Apr 30;11(1):48 [FREE Full text] [doi: 10.1186/1745-6215-11-48] [Medline: 20429947]
147. Fogarty AS, Proudfoot J, Whittle EL, Clarke J, Player MJ, Christensen H, et al. Preliminary evaluation of a brief web and mobile phone intervention for men with depression: men's positive coping strategies and associated depression, resilience, and work and social functioning. JMIR Ment Health 2017 Aug 10;4(3):e33 [FREE Full text] [doi: 10.2196/mental.7769] [Medline: 28798009]
148. Kauer SD, Reid SC, Crooke AH, Khor A, Hearps SJ, Jorm AF, et al. Self-monitoring using mobile phones in the early stages of adolescent depression: randomized controlled trial. J Med Internet Res 2012 Jun;14(3):e67 [FREE Full text] [doi: 10.2196/jmir.1858] [Medline: 22732135]
149. van der Zanden ZR, Kramer J, Gerrits R, Cuijpers P. Effectiveness of an online group course for depression in adolescents and young adults: a randomized trial. J Med Internet Res 2012;14(3):e86 [FREE Full text] [doi: 10.2196/jmir.2033] [Medline: 22677437]
150. Maideen SF, Mohd-Sidik S, Rampal L, Mukhtar F, Ibrahim N, Phang C, et al. A web-based psychoeducational intervention program for depression and anxiety in an adult community in Selangor, Malaysia: protocol of a randomized controlled trial. JMIR Res Protoc 2016 Jun 21;5(2):e112 [FREE Full text] [doi: 10.2196/resprot.4622] [Medline: 27329333]
151. Morris RR, Schueller SM, Picard RW. Efficacy of a web-based, crowdsourced peer-to-peer cognitive reappraisal platform for depression: randomized controlled trial. J Med Internet Res 2015 Mar;17(3):e72 [FREE Full text] [doi: 10.2196/jmir.4167] [Medline: 25835472]
152. Kumagai N, Tajika A, Hasegawa A, Kawanishi N, Horikoshi M, Shimodera S, et al. Predicting recurrence of depression using lifelog data: an explanatory feasibility study with a panel VAR approach. BMC Psychiatry 2019 Dec 11;19(1):391 [FREE Full text] [doi: 10.1186/s12888-019-2382-2] [Medline: 31829206]
153. Migoya-Borja M, Delgado-Gómez D, Carmona-Camacho R, Porras-Segovia A, López-Moriñigo JD, Sánchez-Alonso M, et al. Feasibility of a virtual reality-based psychoeducational tool (VRight) for depressive patients. Cyberpsychol Behav Soc Netw 2020 Apr 01;23(4):246-252. [doi: 10.1089/cyber.2019.0497] [Medline: 32207997]

154. Anttila K, Anttila M, Välimäki M. A web-based adolescent depression support system: feedback and implications for the future. Inform Health Soc Care 2020 Nov 04;45(2):111-129. [doi: 10.1080/17538157.2019.1656206] [Medline: 31684788]
155. Goodwin BC, Ford DE, Hsiung RC, Houston TK, Fogel J, Van Voorhees BW. First, do no harm: referring primary care patients with depression to an internet support group. Telemed J E Health 2018 Jan;24(1):37-44 [FREE Full text] [doi: 10.1089/tmj.2017.0004] [Medline: 28657881]
156. Inkster B, Sarda S, Subramanian V. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study. JMIR Mhealth Uhealth 2018 Nov 23;6(11):e12106 [FREE Full text] [doi: 10.2196/12106] [Medline: 30470676]
157. Soares Teles A, Rocha A, da Silva E Silva FJ, Correia Lopes J, O'Sullivan D, Van de Ven P, et al. Enriching mental health mobile assessment and intervention with situation awareness. Sensors (Basel) 2017 Jan 10;17(1):127 [FREE Full text] [doi: 10.3390/s17010127] [Medline: 28075417]
158. Christensen H, Batterham P, Mackinnon A, Griffiths KM, Kalia Hehir K, Kenardy J, et al. Prevention of generalized anxiety disorder using a web intervention, iChill: randomized controlled trial. J Med Internet Res 2014 Sep 02;16(9):e199 [FREE Full text] [doi: 10.2196/jmir.3507] [Medline: 25270886]
159. Radovic A, Gmelin T, Hua J, Long C, Stein BD, Miller E. Supporting Our Valued Adolescents (SOVA), a social media website for adolescents with depression and/or anxiety: technological feasibility, usability, and acceptability study. JMIR Ment Health 2018 Feb 26;5(1):e17 [FREE Full text] [doi: 10.2196/mental.9441] [Medline: 29483067]
160. Dahne J, Collado A, Lejuez CW, Risco CM, Diaz VA, Coles L, et al. Pilot randomized controlled trial of a Spanish-language behavioral activation mobile app (¡Aptívate!) for the treatment of depressive symptoms among united states Latinx adults with limited English proficiency. J Affect Disord 2019 May 01;250:210-217. [doi: 10.1016/j.jad.2019.03.009] [Medline: 30870770]
161. Segal ZV, Dimidjian S, Beck A, Boggs JM, Vanderkruik R, Metcalf CA, et al. Outcomes of online mindfulness-based cognitive therapy for patients with residual depressive symptoms: a randomized clinical trial. JAMA Psychiatry 2020 Jun 01;77(6):563-573 [FREE Full text] [doi: 10.1001/jamapsychiatry.2019.4693] [Medline: 31995132]
162. Hur J, Kim B, Park D, Choi S. A scenario-based cognitive behavioral therapy mobile app to reduce dysfunctional beliefs in individuals with depression: a randomized controlled trial. Telemed J E Health 2018 Dec;24(9):710-716. [doi: 10.1089/tmj.2017.0214] [Medline: 29323626]
163. Miloff A, Marklund A, Carlbring P. The challenger app for social anxiety disorder: new advances in mobile psychological treatment. Internet Interv 2015 Nov;2(4):382-391. [doi: 10.1016/j.invent.2015.08.001]
164. Geraedts AS, Kleiboer AM, Wiezer NM, van Mechelen W, Cuijpers P. Short-term effects of a web-based guided self-help intervention for employees with depressive symptoms: randomized controlled trial. J Med Internet Res 2014 May 06;16(5):e121 [FREE Full text] [doi: 10.2196/jmir.3185] [Medline: 24800966]
165. Takahashi K, Takada K, Hirao K. Feasibility and preliminary efficacy of a smartphone application intervention for subthreshold depression. Early Interv Psychiatry 2019 Feb;13(1):133-136. [doi: 10.1111/eip.12540] [Medline: 29356332]
166. McCall HC, Richardson CG, Helgadottir FD, Chen FS. Evaluating a web-based social anxiety intervention among university students: randomized controlled trial. J Med Internet Res 2018 Dec 21;20(3):e91 [FREE Full text] [doi: 10.2196/jmir.8630] [Medline: 29563078]
167. Garrido S, Millington C, Cheers D, Boydell K, Schubert E, Meade T, et al. What works and what doesn't work? A systematic review of digital mental health interventions for depression and anxiety in young people. Front Psychiatry 2019 Nov 13;10:759 [FREE Full text] [doi: 10.3389/fpsyt.2019.00759] [Medline: 31798468]
168. Grist R, Porter J, Stallard P. Mental health mobile apps for preadolescents and adolescents: a systematic review. J Med Internet Res 2017 May 25;19(5):e176 [FREE Full text] [doi: 10.2196/jmir.7332] [Medline: 28546138]
169. Clement S, Schauman O, Graham T, Maggioni F, Evans-Lacko S, Bezborodovs N, et al. What is the impact of mental health-related stigma on help-seeking? A systematic review of quantitative and qualitative studies. Psychol Med 2015 Jan;45(1):11-27. [doi: 10.1017/S0033291714000129] [Medline: 24569086]
170. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. BMC Psychiatry 2010;10:113 [FREE Full text] [doi: 10.1186/1471-244X-10-113] [Medline: 21192795]
171. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015 Feb 24;17(2):52-57 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
172. Zarski A, Lehr D, Berking M, Riper H, Cuijpers P, Ebert DD. Adherence to internet-based mobile-supported stress management: a pooled analysis of individual participant data from three randomized controlled trials. J Med Internet Res 2016 Jun 29;18(6):e146 [FREE Full text] [doi: 10.2196/jmir.4493] [Medline: 27357528]
173. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on internet-based mental health interventions — A systematic review. Internet Interven 2014 Oct;1(4):205-215. [doi: 10.1016/j.invent.2014.08.003]
174. Palmqvist B, Carlbring P, Andersson G. Internet-delivered treatments with or without therapist input: does the therapist factor have implications for efficacy and cost? Expert Rev Pharmacoecon Outcomes Res 2007 Jun;7(3):291-297. [doi: 10.1586/14737167.7.3.291] [Medline: 20528315]

175. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
176. Uhlhaas P, Torous J. Digital tools for youth mental health. NPJ Digit Med 2019 Oct 18;2(1):104 [FREE Full text] [doi: 10.1038/s41746-019-0181-2] [Medline: 31646184]
177. Mobile operating system market share worldwide: Nov 2019 - Nov 2020. StatCounter. 2020. URL: https://gs.statcounter.com/os-market-share/mobile/ [accessed 2022-01-10]
178. Carlo AD, Ghomi RH, Renn BN, Areán PA. By the numbers: ratings and utilization of behavioral health mobile applications. NPJ Digit Med 2019 Jun 17;2(1):54 [FREE Full text] [doi: 10.1038/s41746-019-0129-6] [Medline: 31304400]
179. Lin Y, Chen S, Lin P, Tai A, Pan Y, Hsieh C, et al. Assessing user retention of a mobile app: survival analysis. JMIR Mhealth Uhealth 2020 Nov 26;8(11):e16309 [FREE Full text] [doi: 10.2196/16309] [Medline: 33242023]
180. Torous J, Levin ME, Ahern DK, Oser ML. Cognitive behavioral mobile applications: clinical studies, marketplace overview, and research agenda. Cogn Behav Pract 2017 May;24(2):215-225. [doi: 10.1016/j.cbpra.2016.05.007]
181. Wisniewski H, Liu G, Henson P, Vaidyam A, Hajratalli NK, Onnela J, et al. Understanding the quality, effectiveness and attributes of top-rated smartphone health apps. Evid Based Ment Health 2019 Feb;22(1):4-9. [doi: 10.1136/ebmental-2018-300069] [Medline: 30635262]
182. Olff M. Mobile mental health: a challenging research agenda. Eur J Psychotraumatol 2015 May 19;6(1):27882 [FREE Full text] [doi: 10.3402/ejpt.v6.27882] [Medline: 25994025]
183. Huang H, Bashir M. Users' adoption of mental health apps: examining the impact of information cues. JMIR Mhealth Uhealth 2017 Jun 28;5(6):e83 [FREE Full text] [doi: 10.2196/mhealth.6827] [Medline: 28659256]
184. Wei Y, Zheng P, Deng H, Wang X, Li X, Fu H. Design features for improving mobile health intervention user engagement: systematic review and thematic analysis. J Med Internet Res 2020 Dec 09;22(12):e21687 [FREE Full text] [doi: 10.2196/21687] [Medline: 33295292]
185. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018 Aug 05;21(3):116-119. [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
186. Aref-Adib G, McCloud T, Ross J, O'Hanlon P, Appleton V, Rowe S, et al. Factors affecting implementation of digital health interventions for people with psychosis or bipolar disorder, and their family and friends: a systematic review. Lancet Psychiatry 2019 Mar;6(3):257-266. [doi: 10.1016/s2215-0366(18)30302-x]
187. Larsen ME, Huckvale K, Nicholas J, Torous J, Birrell L, Li E, et al. Using science to sell apps: evaluation of mental health app store quality claims. NPJ Digit Med 2019 Mar 22;2(1):18 [FREE Full text] [doi: 10.1038/s41746-019-0093-1] [Medline: 31304366]
188. Nicholas J, Larsen ME, Proudfoot J, Christensen H. Mobile apps for bipolar disorder: a systematic review of features and content quality. J Med Internet Res 2015 Aug;17(8):e198 [FREE Full text] [doi: 10.2196/jmir.4581] [Medline: 26283290]
189. Singh K, Drouin K, Newmark LP, Lee J, Faxvaag A, Rozenblum R, et al. Many mobile health apps target high-need, high-cost populations, but gaps remain. Health Aff (Millwood) 2016 Dec 01;35(12):2310-2318. [doi: 10.1377/hlthaff.2016.0578] [Medline: 27920321]
190. IEEE 1752.1-2021 - IEEE Standard for Open Mobile Health Data--Representation of Metadata, Sleep, and Physical Activity Measures. IEEE Standrads Association. 2021. URL: https://standards.ieee.org/project/1752_1.html [accessed 2021-06-05]
191. HL7 Consumer Mobile Health Application Functional Framework (CMHAFF). HL7 International. URL: https://www.hl7.org/implement/standards/product_brief.cfm?product_id=476 [accessed 2022-01-10]
192. App advisor: an American Psychiatric Association initiative. American Psychiatric Association. URL: https://www.psychiatry.org/psychiatrists/practice/mental-health-apps [accessed 2021-11-01]
193. The Digital Technology Assessment Criteria for Health and Social Care (DTAC). NHSX. 2021. URL: https://www.nhsx.nhs.uk/key-tools-and-info/digital-technology-assessment-criteria-dtac/ [accessed 2022-01-10]
194. Digital Medicine Society. URL: https://www.dimesociety.org/about-us/ [accessed 2021-06-05]
195. Advancing DTx: improving clinical and health economic outcomes. Digital Therapeutics Alliance. URL: https://dtxalliance.org/advancing-dtx/ [accessed 2021-06-05]
196. ISO/PRF TS 82304-2 Health software - Part 2: health and wellness apps - quality and reliability. ISO. URL: https://www.iso.org/standard/78182.html [accessed 2021-05-03]
197. Naden C. One step ahead with mobile apps. ISO. 2020. URL: https://www.iso.org/news/isofocus_141-6.html [accessed 2021-01-11]
198. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
199. Nouri R, Kalhori SR, Ghazisaeedi M, Marchand G, Yasini M. Criteria for assessing the quality of mHealth apps: a systematic review. J Am Med Inform Assoc 2018 Aug 01;25(8):1089-1098 [FREE Full text] [doi: 10.1093/jamia/ocy050] [Medline: 29788283]
200. Dattani S, Ritchie H, Roser M. Mental health. Our World in Data. 2018. URL: https://ourworldindata.org/mental-health [accessed 2022-01-10]

201. Bidargaddi N, Schrader G, Klasnja P, Licinio J, Murphy S. Designing m-Health interventions for precision mental health support. Transl Psychiatry 2020 Jul 07;10(1):222 [FREE Full text] [doi: 10.1038/s41398-020-00895-2] [Medline: 32636358]
202. Kaldy J. Digital therapeutics: health care wired for the future. Senior Care Pharm 2020 Aug 01;35(8):338-344. [doi: 10.4140/tcp.n.2020.338]
203. "Digital Therapeutics: Combining Technology and Evidence-based Medicine to Transform Personalized Patient Care". Digital Therapeutics Alliance. 2018. URL: https://www.dtxalliance.org/wp-content/uploads/2018/09/DTA-Report_DTx-Industry-Foundations.pdf [accessed 2020-12-18]

Abbreviations

CBT: cognitive behavioral therapy
DTx: digital therapeutics
mHealth: mobile health
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO: International Prospective Register of Systematic Reviews
RCT: randomized controlled trial
RWE: real-world evidence

Edited by R Kukafka; submitted 23.01.21; peer-reviewed by M Wolf, E Szigethy; comments to author 02.04.21; revised version received 05.06.21; accepted 08.11.21; published 04.02.22.

Please cite as:
Leong QY, Sridhar S, Blasiak A, Tadeo X, Yeo G, Remus A, Ho D
Characteristics of Mobile Health Platforms for Depression and Anxiety: Content Analysis Through a Systematic Review of the Literature and Systematic Search of Two App Stores
J Med Internet Res 2022;24(2):e27388
URL: https://www.jmir.org/2022/2/e27388
doi:10.2196/27388
PMID:35119370

Original Paper

Web-Based Research Trends on Child and Adolescent Cancer Survivors Over the Last 5 Years: Text Network Analysis and Topic Modeling Study

Hyun-Yong Kim[1*], BSN, MA; Kyung-Ah Kang[2*], DPhil; Suk-Jung Han[2*], DPhil; Jiyoung Chun[2*], DPhil

[1]Logos Health Design Institute, Sahmyook University, Seoul, Republic of Korea

[2]College of Nursing, Sahmyook University, Seoul, Republic of Korea

[*]all authors contributed equally

Corresponding Author:
Kyung-Ah Kang, DPhil
College of Nursing
Sahmyook University
815, Hwarang-ro, Nowon-gu
Seoul, 01795
Republic of Korea
Phone: 82 10 4709 4731
Fax: 82 2 3399 1594
Email: kangka@syu.ac.kr

Abstract

Background: Being diagnosed with cancer during childhood or adolescence can disrupt important periods in an individual's physical, psychosocial, and spiritual development and potentially reduce the quality of life (QOL) after treatment. Research is urgently required to improve the QOL for child and adolescent cancer survivors, and it is necessary to analyze the trends in prior research reported in international academic journals to identify knowledge structures.

Objective: This study aims to identify the main keywords based on network centrality, subgroups (clusters) of keyword networks by using a cohesion analysis method, and the main theme of child and adolescent cancer survivor–related research abstracts through topic modeling. This study also aims to label the subgroups by comparing the results of the cohesion and topic modeling.

Methods: A text network analysis method and topic modeling were used to explore the main trends in child and adolescent cancer survivor research by structuring a network of keyword (semantic morphemes) co-occurrence in the abstracts of articles published in 5 major web-based databases from 2016 to 2020. A total of 1677 child and adolescent cancer survivor–related studies were used for data analyses. Data selection, processing, and analyses were also conducted.

Results: The top 5 keywords in terms of degree and eigenvector centrality were *risk*, *control interval*, *radiation*, *childhood cancer treatment*, and *diagnosis*. Of the 1677 studies used for data analyses, cluster 1 included 780 (46.51%) documents under *risk management*, cluster 2 contained 557 (33.21%) articles under *health-related QOL and supportive care*, and cluster 3 consisted of 340 (20.27%) studies under *cancer treatment* and *complications*.

Conclusions: This study is significant in that it confirms the knowledge structure based on the main keywords and cross-disciplinary trends in child and adolescent cancer survivor research published in the last 5 years worldwide. The primary goal of child and adolescent cancer survivor research is to prevent and manage the various aspects of the problems encountered during the transition to a normal life and to improve the overall QOL. To this end, it is necessary to further revitalize the study of the multidisciplinary team approach for the promotion of age-specific health behaviors and the development of intervention strategies with increased feasibility for child and adolescent cancer survivors.

(J Med Internet Res 2022;24(2):e32309) doi:10.2196/32309

KEYWORDS

text network analysis; topic modeling; cancer survivors; child; adolescent; research trends; knowledge structures

Introduction

Background

More than 300,000 children worldwide are diagnosed with cancer every year; every 3 minutes, a family somewhere in the world receives the unexpected and shattering news that their child has been diagnosed with cancer [1]. According to a 2017 survey of major causes of death across all races and sexes in the United States, malignant neoplasm was the leading cause of disease-related death in children and adolescents [2]. Approximately 137 million children are predicted to be diagnosed with cancer worldwide between 2020 and 2050 [3].

With significant advances in diagnosis and increase in multimodal treatment, childhood cancer survival rates have improved significantly in recent decades [4]. In the case of 5-year survivors of childhood cancer, the 30-year cumulative survival rate is 81.9% (95% CI 81.1-82.7) [5]. However, most child and adolescent cancer survivors experience additional problems and health effects after cancer treatment. More than 60% of survivors report with at least one chronic disease, and more than one-third report experiencing a reduced quality of life (QOL) with at least 2 complications [6,7].

Late complications experienced by child and adolescent cancer survivors may be psychosocial and behavioral, such as depression, anxiety, and risky health behaviors, or physical, such as cardiovascular disease, secondary cancer, and hormonal and immune deficiencies [8,9]. Even after full recovery, many child and adolescent cancer survivors have a high risk of death and morbidity resulting from secondary infections [10,11]. In other words, being diagnosed with cancer at this age can disrupt key periods of physical, social, and psychological development and potentially reduce QOL after treatment [12,13]. Therefore, significant effort is required to improve the QOL of child and adolescent cancer survivors.

More targeted research is needed to improve the QOL of child and adolescent cancer survivors; it is necessary to analyze previous research and organize the research results. A systematic literature review and meta-analysis, which are common methods used to intensively analyze intervention effects in the field of interest, are based on specific research questions and analyzed for a limited number of relevant papers [14]. Although it is possible to investigate the outcome of a topic of interest in this way, there is a limitation in that the overall knowledge structure of the context in which the major research concept and topic are studied or the type of connections between the main concepts cannot be known.

Exploiting a text network analysis to study research trends provides useful information with new ideas for future research [15,16]. These results can be used to form a body of knowledge by quantitatively analyzing a large number of papers and qualitatively interpreting the discovered knowledge structure to identify a certain pattern in the scattered data [17]. In addition, because numerous related papers are collected and used for analysis, this method has a significant quantitative scientific basis in terms of the research results and in identifying core research topics in academic fields and its relevance to various subtopics. In particular, it is possible to intuitively explore the knowledge structure by visualizing such research results [18,19].

Through text network analysis, it is possible to check the connectivity (links) between the words included in the documents (in our case, the words in the abstracts) and the connectivity group (community) of keywords based on their links. In addition, if a topic modeling analysis, which has the advantage of identifying topics representing the community based on the probability distribution of words included in the document, is applied simultaneously within a text network analysis, then trends based on the core topic of research in the field of interest can be clearly identified. Although the probability distribution of words in a document does not have any intuitive meaning, researchers can interpret the meaning of a specific topic, and thus, the extracted topic can be used as important information representing the document [20,21].

In the field of adult cancer and cancer survivor care, the number of studies applying a text network analysis to confirm the knowledge structure and overall research trends in the field of interest have increased over the past 10 years [16,22-25]. However, few studies that were published in 2017 applied a network analysis method to child and adolescent cancer survivors [15]. Because information technology has been undergoing rapid advancements, it is important that research on child and adolescent cancer survivors is shared all over the world.

To improve the QOL of child and adolescent cancer survivors based on the latest information, it is necessary to comprehensively identify research trends related to child and adolescent cancer survivors using text network analysis on studies published in internationally recognized academic databases. Therefore, in this study, using a text network analysis method and topic modeling, the main research concepts of previous studies on child and adolescent cancer survivors published in international academic journals in the last 5 years were reviewed, and research trends were identified.

Objectives

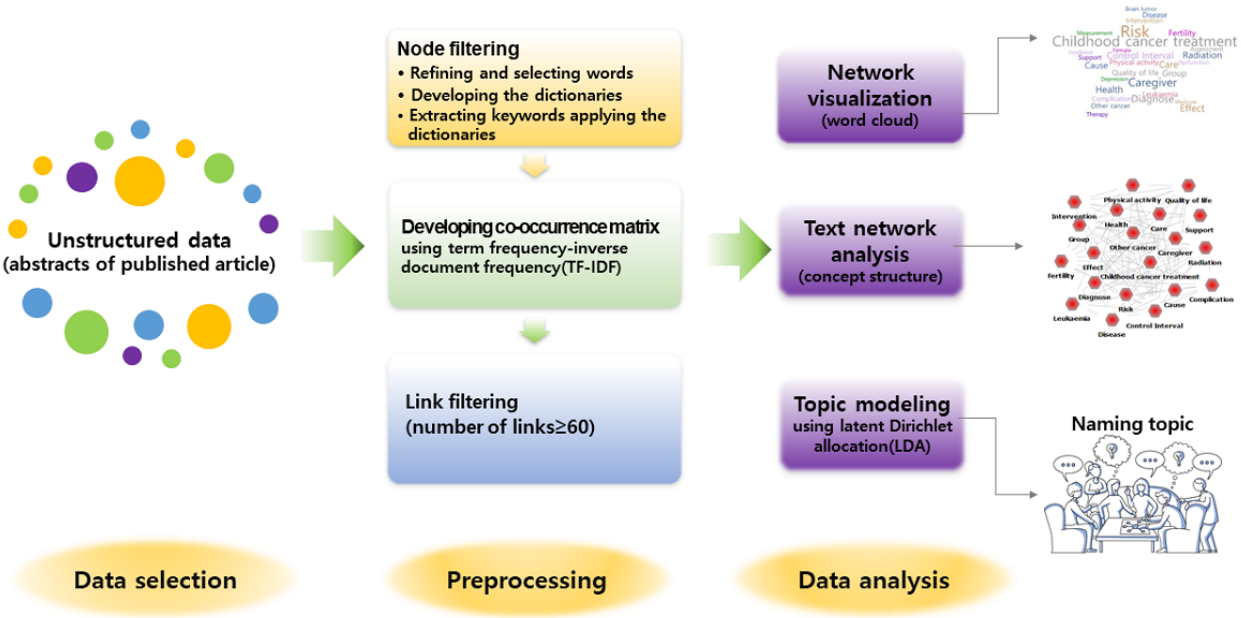
The specific goals of this study that analyze the abstracts of child and adolescent cancer survivor–related research papers are as follows: identifying the main keywords based on network centrality, finding the subgroups of a network of keywords using a cohesion analysis, clustering the main theme of child and adolescent cancer survivor–related research abstracts through topic modeling, and comparing and analyzing the results of cohesion and topic modeling.

Methods

Study Design

This study used text network analysis and topic modeling to explore the main research trends of child and adolescent cancer survivors by structuring a word co-occurrence network among the keywords (semantic morphemes) in the abstracts of articles published in 5 major web-based databases between 2016 and 2020. The process used is illustrated in Figure 1.

Figure 1. The process of this study.



Ethical Considerations

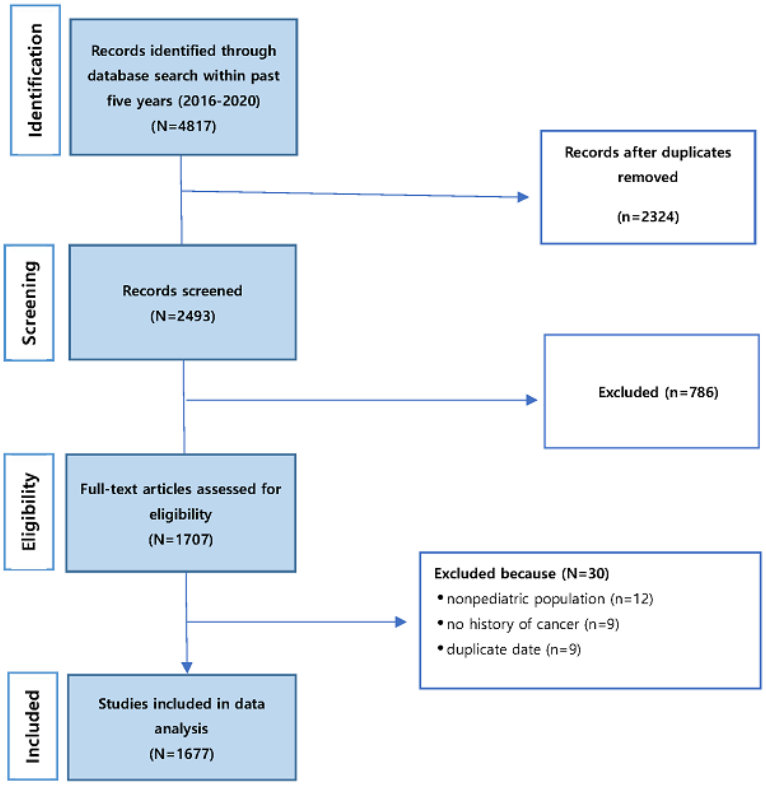
This study was approved by the Institutional Review Board of Sahmyook University (IRB No. 2020116HR). The data were collected from published articles in web-based databases, and no harm or risk was posed by the research to the participants. In addition, the collected data will not be used for any purpose other than for this study.

Data Selection

Keyword Selection

The first search to retrieve published articles from all time periods in 5 major databases (PubMed, CINAHL, Cochrane, Embase, and PsycInfo) was conducted in January 2021. The main keywords used included *Adolescent* [Mesh] OR *Child* [Mesh] OR *Pediatrics* [Mesh] and *Cancer Survivors* [Mesh] along with *Childhood cancer survivors*, *Adolescent cancer survivors*, and *Pediatric cancer survivors*. Only articles written in English during the 5-year period of 2016-2020 were selected. The article identification and selection process used the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram protocol (Figure 2) to retrieve 4817 articles during the initial search; 51.75% (2493/4817) articles were retained after eliminating duplicates (2324/4817, 48.25%). Each title was reviewed in accordance with this study's topic of interest, and 31.52% (786/2493) irrelevant studies were excluded. In all, 68.47% (1707/2493) of full-text articles were assessed for eligibility. Of the 1707 articles assessed, 30 (1.76%) that included a nonpediatric population (12/30, 40%), no history of cancer (9/30, 30%), and duplicate data (9/30, 30%) were excluded. Finally, 67.27% (1677/2493) abstracts were collected after reviewing their adequacy, and a network analysis was conducted on the data.

Figure 2. A PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.



Collection of Abstracts

A total of 2 researchers (HYK and KAK) of this study reviewed the titles of the retrieved data to identify the criteria for selection and exclusion. Moreover, 2 other researchers (SJH and JC) of this study repeated the former review process. A total of 1677 studies were selected for the final data analyses.

Preprocessing

Overview

The final selected abstracts were categorized to generate compatible unstructured data that could be accessed by NetMiner (Cyram) version 4.0. NetMiner is a validated program that conducts text network analysis by extracting a keyword from the abstracts and applying automatic filtration of the parts of speech (eg, nouns, verbs, and adjectives), generating a word co-occurrence matrix, and visualizing the analysis results [26].

First, to create compatible unstructured data, we categorized the retrieved studies based on publication year, author, title, abstract, and keywords using Microsoft Excel 2016. The raw data were then imported into NetMiner, and 10,786 words were identified as the main set of nodes.

Node Filtering

Word Refining and Selection

In all, 2 researchers of this study conducted a brief review of the first main node set of 10,786 words, regardless of the occurrence frequency. There is no scientific protocol available for the cut-off standard; however, it is widely accepted that words that only appear once can be excluded. During the review process, most of the words with a frequency of 1-4 showed inaccurate spellings or errors, irrelevance with respect to the main concept of the study, and general verbs or adjectives that were considered far removed from the actual content. Therefore, both researchers reached a consensus to exclude words with a frequency of <5, thus leaving 2755 words as a result.

Development of Dictionaries

For meaningful text network analysis, refining the semantic morphemes is important [17,26]. For the further filtering process, 4 researchers conducted a detailed review of each term and created a filtering dictionary. A total of 3 types of filtering dictionaries were created during this review. First, the thesaurus dictionary contained the identification of keywords that had the same definition but were listed in different ways (eg, *QOL*, *quality of life*, and *Quality of Life*). An exclusion dictionary was created to eliminate words that were not significantly related to the study content (eg, *apple*, *Netherlands*, and *cars*), and terms that were composed of multiple words representing a single definition (eg, *physical activity* as a single term, not *physical* and *activity* separately) were listed in the defined dictionary.

Keywords Extracted by Applying Dictionaries

The dictionaries were applied using the NetMiner software while importing unstructured data. With a network structure in which people in a social network analysis are nodes (points) and relationships among people are links (lines), a text network

analysis helps us understand the relationship among words comprehensively and intuitively by providing a sociogram in which words are applied as nodes, and co-occurrence relationships are applied as links [18]. This filtering process was repeated until the main node set of words was clearly noticeable without any errors (eg, uncapitalized terms, undefined terms, or unorganized synonyms). When the first filtering process was completed, 217 words remained in the main node set.

Development of Co-occurrence Matrix for Keywords (From 2-Mode to 1-Mode Network)

In this study, an abstract was considered a single document unit; each term represented a single node unit, and the initial word-document network was created (2-mode). Then, to process the comembership between terms, a co-occurrence word network (1-mode) was generated using the recoded weight value of term frequency-inverse document frequency (TF-IDF). The TF-IDF value reflects both the importance and frequency of a term occurring in a set of documents; therefore, it can be considered as an appropriate value for conducting a keyword network analysis [17,18,26]. An inner-product-proximity measure was adopted during the mode-conversion process.

Link Filtering

The generated 1-mode network consisted of 10,826 links. An excessive number of links in a network must be filtered to clearly view the results for network visualization. The cut-off standard for link reduction was determined by conducting a link-reduction simulation. A link frequency of ≥60 was chosen for the reduction, and 378 links were finally included in the data analysis.

Data Analysis

NetMiner version 4.0. was used for data analyses.

Network Visualization

Network visualization is the styling of nodes and links such that the connecting structure can be intuitively expressed [17,20]. A total of 3 simulations (with a link frequency of ≥55, ≥60, and ≥65) were conducted to visualize the map of the 1-mode network, and a consensus among the authors was reached to adopt a link frequency of ≥60, which best reflects the major trends in childhood cancer research. We applied a word cloud to visualize the network.

Text Network Analysis

Centrality is a useful index employed to determine the node (keyword) that lies at the center of the network, and the rank is determined based on its relativity rather than the absolute size. A node with a high centrality value can be considered as the core-acting node in a network [17,18,27]. This study adopted the following two approaches: degree and eigenvector centrality.

Degree centrality measures the number of connections carried by a single node, which indicates that a node with a higher degree centrality value and a higher number of connections within a network can be determined as a core keyword. The concept of eigenvector centrality is somewhat similar; however, rather than determining the direct connections of a node, it overviews the entire network to discover the most influential keyword by counting its surrounding neighbors. The standardized value of centrality lies between 0 and 1 and is used to determine the core-acting keywords in the theme of this study [17,26].

A cohesion analysis was conducted to identify the subtopic groups and the division of communities belonging to the co-occurrence word network [17,28]. In the first step, the giant component was extracted, and a community analysis was conducted. The largest modularity (best cut) is regarded as the most optimized value, with values of 1.25-2.75 indicating normal values and 2.75-3.5 indicating good values [17,28]. A higher modularity with a positive value means that the community is significantly divided; this link density within the group is high and between groups is low [17,18,28].

When visualizing the results of the network analysis, it is difficult to interpret the connections between many nodes and links because of their complexity. In this study, a pathfinder network, which is abbreviated such that only important links are left for each node, was used as a graph-drawing algorithm for centrality and cohesion analysis.

Topic Modeling

The topic modeling method, which is a technique used in text mining, assumes that a document is a set of words and that each document contains multiple topics with a specific probability distribution [19-21,27]. For topic modeling, this study applied latent Dirichlet allocation (LDA), a classic type of machine learning method, to discover topic keywords that best represented the core value in the collection of documents [19,21,26]. In this study, a topic modeling analysis applying LDA was used to validate and provide evidence for a cohesion analysis which was the categorization of themes concluded by the authors. LDA was conducted using NetMiner 4.0. The values of the parameter settings were α=.01 and β=.01, and the number of iterations was 1000. The adequate distribution of topics was set at 7, showing a clear distinction between all topics. Probable keywords for each topic were chained to refer to the categorization results of the cohesion analysis.

Results

Characteristics of Included Research

In the last 5 years, an average of 335.4 articles per year were published on child and adolescent cancer survivors (SD 65.24). The number of studies was 252 in 2020, 429 in 2019, 303 in 2018, 349 in 2017, and 344 in 2016. The year in which the most number of child and adolescent cancer survivor–related articles were published was 2019, whereas the most recent year, 2020, had the lowest number of publications.

Text Network Analysis

Degree Rank of 2-Mode (Word Documents), Degree, and Eigenvector Centrality of Keywords

Multimedia Appendix 1 presents the top 20–ranked keywords in the 2-mode network (word documents), degree centrality, and eigenvector centrality. Keywords that appeared most

frequently in the documents included childhood cancer treatment (degree=925), risk (degree=748), effect (degree=681), diagnosis (degree=652), and cause (degree=578). In contrast, the top 5 keywords in terms of degree centrality were *risk*, *control interval*, *childhood cancer treatment*, *diagnosis*, and *radiation*. The degree centrality index was 26.606%, with a mean of 0.018 (SD 0.04). The index of the centrality ranges from 0% to 100%, indicating that the links were fairly distributed within the network when the percentage is closer to 0% [20]. The top 5 keywords in the eigenvector centrality were *risk*, *control interval*, *radiation*, *childhood cancer treatment*, and *respiratory rate*. The mean was 0.025 (SD 0.064. The trends in the theme of the keywords, in both degree and eigenvector centrality rank, revealed their similarity because 75% (15/20) of the keywords overlapped with each other.

Cohesion With Network Visualization

For cohesion analysis, the giant component was extracted from the 1-mode network, which represented the relationship between a set of keywords (main node vs main node) and documents (abstracts). The component included 90 keywords from the network consisting of link frequencies >60 based on the TF-IDF recoded value. A component community analysis, using the Blondel processing method, divided the groups into 3 clusters with a maximum modularity value of 0.258. Textbox 1 lists the clusters divided according to their connection strengths. Cluster 1 contains 34 keywords, including *risk*, *control interval*, *female*, *condition*, and *cancer center*. Cluster 2 contains 31 keywords, including *caregiver*, *effect*, *care*, *health*, and *group*. Cluster 3 contains 25 keywords, including *childhood cancer treatment*, *diagnosis*, *radiation*, *cause*, and *leukemia*. Multimedia Appendix 2 shows the top 10 keywords of each cluster and map of the communities.

Textbox 1. Results of cohesion (clusters; maximum modularity=0.258) and included keywords.

Cluster 1 (n=34): risk management

- *Behavior, BMI, Cancer Center, cardiovascular Disease, Cohort, comorbidity, Comparison, Condition, Control Interval, Death, Deficiency, Diabetes, Engagement, Ethnicity, Exposure, Female, Gonadal, Hormone, Hospitalization, human papillomavirus, Insurance, Knowledge, Male, Malignancy, Mortality, Peer, Perception, Prognosis, Pulmonary, respiratory rate, Risk, Sex, Sibling, Surveillance.*

Cluster 2 (n=31): health-related quality of life and supportive care

- *Anxiety, Burden, Care, Caregiver, Central control station, Concern, Counseling, Depression, Education, Effect, Fatigue, Fertility, Group, Health, health care professional, Information, Intervention, Measurement, Nutrition, Physical activity, Physician, Posttraumatic growth, quality of life, Resilience, School, Stress, Support, Survivorship, Symptom, Transition, well-being.*

Cluster 3 (n=25): cancer treatment and complication

- *5 y, Assessment, BMD, Brain tumor, Cardiac, Cause, Central Nervous System, Chemotherapy, Childhood cancer treatment, Complication, Diagnose, Disease, Dysfunction, Experience, Incidence, leukemia, Lymphoma, magnetic resonance imaging, Medicine, Operation, Other cancer, Radiation, Regimen, Therapy, Transplant.*

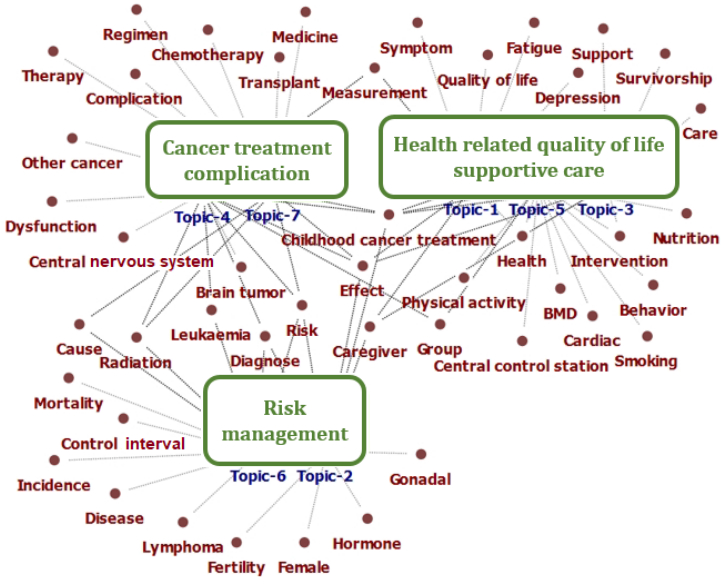
Topic Modeling

The keywords in the results of topic modeling, when applying LDA, feature the probability of being the best representation for the classified documents. In this study, 7 topics were determined to be appropriate matches in accordance with the results of the cohesion analysis. Multimedia Appendix 3 presents the keywords representing each topic based on a probability distribution with the number of documents included for each topic. Among the keywords representing the documents belonging to each topic, the first keyword has the highest probability of representing each topic, and the probability decreases with each keyword [28]. In this study, only the first and second keywords were considered for the analysis. Of the 1677 documents, topic 1 consisted of 265 (15.8%) documents, and the most probable keywords representing the documents belonging to this topic in order were *care* or *caregiver*. Topic 2 consisted of the lowest number of documents at 6.97% (117/1677), with *fertility* or *female* as representative keywords. Topic 3 had the largest number of documents at 21.36% (358/1677), with *care* or *caregiver* as the representative keyword. Topic 4 included 12.52% (210/1677) documents and was represented by the keywords *brain tumor* or *radiation*. Topic 5 consisted of 9.36% (157/1677) documents, with keywords including *physical activity* and *health*. Topic 6 consisted of 13.30% (223/1677) documents, with *risk* or *control interval* as representative keywords. Finally, Topic 7 contained the second largest number of documents at 20.69% (347/1677), with *complication* and *radiation* as representative keywords.

Comparison of Results of Cohesion With Topic Modeling and Topic Naming

On the basis of the keyword trends in each community, analyzed through cohesion, the authors reviewed the similarities among keywords, included in each cluster, in the results of the cohesion as well as each topic with representative keywords in the results of the topic modeling. As shown in Textbox 1 and Multimedia Appendix 3, comparing topics 2 and 6, we reached a consensus to name cluster 1 as *risk management*. Cluster 2 with topics 1, 3, and 5 was labeled *health-related QOL and supportive care*, and cluster 3 with topics 4 and 7 was labeled *cancer treatment and complication*. Figure 3 presents a clearer comparison of the keywords surrounding each categorized topic based on cohesion analysis and topic modeling.

Figure 3. Comparison map of results between cohesion analysis and topic modeling.



Discussion

Principal Findings

In this study, child and adolescent cancer survivor–related research trends over the past 5 years were analyzed by reviewing papers published in all academic fields.

The main findings of this study were that the top 5 keywords in terms of degree and eigenvector centrality are risk, control interval, radiation, childhood cancer treatment, and diagnosis. Of 1677 documents, cluster 1 included 780 (46.51%) documents under *risk management*, cluster 2 contained 557 (33.21%) articles under *health-related QOL and supportive care*, and cluster 3 consisted of 340 (20.27%) studies under *cancer treatment* and *complications*.

Among the documents included in the 3 subgroups, the knowledge structure and research trends were described by comparing the results of the meta-analysis that presented research results regarding the concerned topics. The giant component from the centrality analysis in this study showed 3 subgroups based on the cohesion analysis. To name the cluster of subgroups more accurately, we compared the results of the cohesion analysis and topic modeling. Cohesion analysis has the advantage of showing the degree of cohesion between keywords belonging to the relevant community [17,28]. However, there is a limitation in that it does not have guidelines based on mathematical calculations to name a given community. Recently, topic modeling, in which keywords that can represent documents belonging to each topic are presented based on probability values, has frequently been used to supplement the limitations of cohesion analysis [20,21]. On the basis of keyword trends in each community with cohesion analysis, the authors reviewed the similarity among keywords included in each cluster in the results of cohesion analysis and each topic with representative keywords in the results of topic modeling. In addition, keywords representing each topic should be mutually exclusive [26]. In the topic modeling results of this study, the first and second keywords representing each topic did not overlap and were mutually exclusive. Therefore, the cohesion analysis and topic modeling results indicate that the clusters of keywords representing recent research trends were reasonably named.

In cluster 1, labeled *risk management*, keywords related to periodic risk management and various risks that child and adolescent cancer survivors could experience while living as survivors after cancer treatment were included. A total of 10 meta-analysis studies corresponding to cluster 1 were identified. In all, 6 meta-papers [29-34] regarding long-term risk and management needs were published, and smoking, binge drinking, and drug use were reported as vulnerable issues for child and adolescent cancer survivors. In particular, the importance of periodic surveillance for late risk management such as cardiac toxicity was reported in a study related to surveillance guidelines among 4 meta-papers [35-38] related to survivorship in child and adolescent cancer survivors. On the basis of this analysis, it is necessary to conduct various studies related to education, information provisioning, health care–provider training, and social support and care system establishment related to periodic surveillance systems for potential risk management.

In cluster 2, labeled "health-related QOL and supportive care," keywords related to the various experiences of child and adolescent cancer survivors' care, and QOL required for survivorship were included. A total of 30 articles were identified in the meta-analysis studies corresponding to cluster 2. Most meta-studies on the development and effect of interventions related to physical health behaviors, such as exercise, nutrition, health promotion behavior, and obesity prevention, were found in 13 articles [39-51]. The results of this study related to keywords such as *effect*, *health*, *physical activity*, *intervention*, and *nutrition* were also confirmed. Specifically, studies proposed that future research should focus on providing evidence of the efficiency and feasibility of interventions that use web-based technologies (such as @TheTable e-cookbook) to facilitate

remote intervention delivery and peer support as the most effective means of promoting changes in health behavior [39-41,49,50,52]. We identified *anxiety*, *burden*, *depression*, *resilience*, *stress*, *counseling*, *measurement*, and *post-traumatic growth*. In 7 meta-analyses [53-59], child and adolescent cancer survivors were highly likely to experience mood or affective disorders [56], and brain tumor in child and adolescent cancer survivors was significantly associated with anxiety, depression, and health-related QOL, highlighting the importance of psychosocial screening. Furthermore, in 1 systematic review and meta-study [54], dealing with the keyword *post-traumatic growth* identified in this study, the need for targeted social support, clinical intervention, and education to facilitate posttraumatic growth was suggested through an analysis of 18 studies. Other keywords related to academics, school, and social adaptation were *education*, *group*, *school*, and *transition*, with 6 meta-studies [60-65] related to these keywords. In particular, the results of 26 studies were analyzed in a meta-study conducted by Saatci et al [65], and compared with controls there were significant differences in educational attainment among survivors according to country and culture. It has been reported that the support and participation of a cross-professional group, such as that of clinicians, teachers, and policymakers, is needed, and the problem of academic and school adjustment for child and adolescent cancer survivors involves an extremely critical developmental issue and an important global care problem that should not be overlooked. A total of 3 studies [66-68] addressed the fertility issues among child and adolescent cancer survivors. The most commonly reported barriers were a lack of patient educational materials regarding oncofertility psychosocial support and staff training [67], whereas another study attempted to realize a qualitative meta-synthesis [69] by synthesizing 51 research results into 5 themes (ie, experiencing difficulties related to cancer, fluctuating realities, coping strategies, new roles and responsibilities of the child, and practical resources to enable the managing of cancer). Compared with the keyword-analysis results confirmed for cluster 2 of this study, it is necessary to continuously conduct research on adaptation and improvement of QOL for children with cancer in terms of their physical, psychological, social, future occupational, and reproductive problems.

Cluster 3 of the cohesive analysis was labeled *cancer treatment and complication* because the main keywords were matched with topics 4 and 7 of the topic modeling results. Among the relevant papers, 11 meta-analysis studies [48,70-79] were identified, and the correlation with keywords included in cluster 3 of this study was confirmed. Keywords related to major carcinomas were *leukemia*, *brain tumor*, and *lymphoma*, and for other cancer types, *other cancer* were designated as representative words. Words related to *childhood cancer treatment* and *diagnosis* were the main keywords. Meta-studies were conducted on the subject of complications and therapeutic management of the endocrine system [78], nervous system [71,77,79], heart [48,72-74], tinnitus [75], and infection [76]. In particular, meta-analyses related to cardiac toxicity were predominant in 4 articles [48,72-74]. Most of the studies conducted in the medical field were identified as the main characteristics of the studies that fall into cluster 3, and the characteristics of interdisciplinary research topic keywords were confirmed.

Implications

The purpose of this study was to identify the knowledge structure based on major research trends and core research topics related to all child and adolescent cancer survivors. The 3 categories (*risk management, health-related QOL and supportive care*, and *cancer treatment* and *complications*) imply the current research trends. The results of this study are significant in that they can suggest practical research directions for the specific needs regarding QOL of child and adolescent cancer survivors in all academic fields.

In terms of the significance of the research methodology, this is the first study to use a text network analysis on child and adolescent cancer survivor–related topics and categorize interdisciplinary research trends. It has the advantage of increasing the validity of research trends as a quantitative method, based on a higher probability, by comparing a cohesion analysis with topic modeling.

Limitations

This study included data from only the last 5 years. The limitation is that the trends in child and adolescent cancer survivor–related research conducted based on the past 5- or 10-year periods cannot be compared or analyzed. The results of this study, focusing on the meta-analysis results included in each cluster, were analyzed and compared with 1677 studies in this research area. Therefore, potential limitations could exist, as some topics have been addressed infrequently in some studies; therefore, these might have been overlooked. Finally, studies in the genetic and biochemical fields were excluded.

Conclusions

This study is significant, in that it confirmed the knowledge structure based on the main keywords and cross-disciplinary research trends related to child and adolescent cancer survivors published in the last 5 years worldwide. To more accurately name the knowledge structure on a quantitative basis, the results of coherence analysis and topic modeling were compared and analyzed based on the giant component confirmed through a centrality analysis. In both cluster 1, *risk management*, and cluster 2, *health-related QOL and supportive care*, many research results for multidisciplinary teams have also been published.

Regarding cluster 3—*cancer treatment and complications*—papers published in the medical field prevailed, confirming the identity of interdisciplinary research topics. Improving QOL is the primary goal of child and adolescent cancer survivors, preventing and managing various aspects of problems encountered during the transition to normal life. To this end, it is necessary to further revitalize the studies through a multidisciplinary team approach to promote age-specific health behaviors and develop intervention strategies with increased feasibility and educational effects for child and adolescent cancer survivors.

Acknowledgments

This study was supported by a National Research Foundation of Korea grant funded by the Korean government (Ministry of Science and Information and Communication Technologies; Grant #2020R1A2C1100912).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Top 20 keywords by degree, degree centrality, and eigenvector centrality.
[PNG File , 53 KB - jmir_v24i2e32309_app1.png]

Multimedia Appendix 2
Top keywords and maps of clusters.
[PNG File , 386 KB - jmir_v24i2e32309_app2.png]

Multimedia Appendix 3
Results of topic modeling (using latent Dirichlet allocation; N=1677).
[PNG File , 63 KB - jmir_v24i2e32309_app3.png]

References

1. NIEHS/EPA Children's environmental health and disease prevention research centers impact report: protecting children's health where they live, learn, and play. EPA Publication No. EPA/600/R-17/407. U.S. Environmental Protection Agency. 2017. URL: https://www.epa.gov/sites/production/files/2017-10/documents/niehs_epa_childrens_centers_impact_report_2017_0.pdf?pdf=chidrens-center-report [accessed 2021-05-16]
2. Heron M. Deaths: leading causes for 2017. Natl Vital Stat Rep 2019 Jun;68(6):1-77 [FREE Full text] [Medline: 32501203]
3. Atun R, Bhakta N, Denburg A, Frazier AL, Friedrich P, Gupta S, et al. Sustainable care for children with cancer: a Lancet Oncology Commission. Lancet Oncol 2020 Apr;21(4):185-224. [doi: 10.1016/S1470-2045(20)30022-X] [Medline: 32240612]
4. Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RG, Barzi A, et al. Colorectal cancer statistics, 2017. CA Cancer J Clin 2017 Dec 06;67(3):177-193 [FREE Full text] [doi: 10.3322/caac.21395] [Medline: 28248415]
5. Berkman AM, Livingston JA, Merriman K, Hildebrandt M, Wang J, Dibaj S, et al. Long-term survival among 5-year survivors of adolescent and young adult cancer. Cancer 2020 Aug 15;126(16):3708-3718. [doi: 10.1002/cncr.33003] [Medline: 32484922]
6. Oeffinger KC, Mertens AC, Sklar CA, Kawashima T, Hudson MM, Meadows AT, Childhood Cancer Survivor Study. Chronic health conditions in adult survivors of childhood cancer. N Engl J Med 2006 Oct 12;355(15):1572-1582. [doi: 10.1056/NEJMsa060185] [Medline: 17035650]
7. Robison LL, Hudson MM. Survivors of childhood and adolescent cancer: life-long risks and responsibilities. Nat Rev Cancer 2014 Jan;14(1):61-70 [FREE Full text] [doi: 10.1038/nrc3634] [Medline: 24304873]
8. Friend AJ, Feltbower RG, Hughes EJ, Dye KP, Glaser AW. Mental health of long-term survivors of childhood and young adult cancer: a systematic review. Int J Cancer 2018 Sep 15;143(6):1279-1286 [FREE Full text] [doi: 10.1002/ijc.31337] [Medline: 29468674]
9. Landier W, Skinner R, Wallace WH, Hjorth L, Mulder RL, Wong FL, et al. Surveillance for late effects in childhood cancer survivors. J Clin Oncol 2018 Jul 20;36(21):2216-2222 [FREE Full text] [doi: 10.1200/JCO.2017.77.0180] [Medline: 29874139]
10. Armstrong GT, Chen Y, Yasui Y, Leisenring W, Gibson TM, Mertens AC, et al. Reduction in late mortality among 5-year survivors of childhood cancer. N Engl J Med 2016 Mar 03;374(9):833-842 [FREE Full text] [doi: 10.1056/NEJMoa1510795] [Medline: 26761625]
11. Ness KK, Howell CR, Bjornard KL. Frailty and quality of life in adult survivors of childhood cancer. Expert Rev Qual Life Cancer Care 2017;2(2):79-85 [FREE Full text] [doi: 10.1080/23809000.2017.1300507] [Medline: 28451645]
12. Smith AW, Parsons HM, Kent EE, Bellizzi K, Zebrack BJ, Keel G, AYA HOPE Study Collaborative Group. Unmet support service needs and health-related quality of life among adolescents and young adults with cancer: the AYA HOPE study. Front Oncol 2013;3:75 [FREE Full text] [doi: 10.3389/fonc.2013.00075] [Medline: 23580328]
13. Bright CJ, Reulen RC, Winter DL, Stark DP, McCabe MG, Edgar AB, et al. Risk of subsequent primary neoplasms in survivors of adolescent and young adult cancer (Teenage and Young Adult Cancer Survivor Study): a population-based, cohort study. Lancet Oncol 2019 Apr;20(4):531-545 [FREE Full text] [doi: 10.1016/S1470-2045(18)30903-3] [Medline: 30797674]
14. Verstraete M. Value and limitation of meta-analysis. Pathophysiol Haemost Thromb 2002;32(5-6):278-281 [FREE Full text] [doi: 10.1159/000073581] [Medline: 13679657]

15. Koltai K, Walsh C, Jones B, Berkelaar BL. Applying social network analysis to identify the social support needs of adolescent and young adult cancer patients and survivors. J Adolesc Young Adult Oncol 2018 Apr;7(2):181-186. [doi: 10.1089/jayao.2017.0058] [Medline: 29106316]
16. Jones EC, Storksdieck M. Recent research on the social network concept and cancer. Curr Opin Support Palliat Care 2019 Sep;13(3):225-237. [doi: 10.1097/SPC.0000000000000442] [Medline: 31246595]
17. Lee S. Network Analysis Methodology. Seoul: Nonhyung; 2012:370.
18. Wasserman S, Faust K. Social Network Analysis: Methods and Applications. Cambridge, United Kingdom: Cambridge University Press; 1994.
19. Gerlach M, Peixoto TP, Altmann EG. A network approach to topic models. Sci Adv 2018 Jul;4(7):eaaq1360 [FREE Full text] [doi: 10.1126/sciadv.aaq1360] [Medline: 30035215]
20. Blei DM, Lafferty JD. Topic Models. In: Text Mining: Classification, Clustering, and Applications. 1st Edition ed. New York: Chapman and Hall/CRC; 2009.
21. Blei DM. Probabilistic topic models. Commun ACM 2012 Apr 01;55(4):77-84. [doi: 10.1145/2133806.2133826]
22. Kwon SY, Bae KR. [A study on the knowledge structure of cancer survivors based on social network analysis]. J Korean Acad Nurs 2016 Feb;46(1):50-58. [doi: 10.4040/jkan.2016.46.1.50] [Medline: 26963414]
23. Henneghan A, Wright ML, Bourne G, Sales AC. A cross-sectional exploration of cytokine-symptom networks in breast cancer survivors using network analysis. Can J Nurs Res 2020 Jun 01:844562120927535. [doi: 10.1177/0844562120927535] [Medline: 32482100]
24. Kim K, Lee K. Identification of the knowledge structure of cancer survivors' return to work and quality of life: a text network analysis. Int J Environ Res Public Health 2020 Dec 14;17(24) [FREE Full text] [doi: 10.3390/ijerph17249368] [Medline: 33327622]
25. Yang K, Kim J, Chun M, Ahn MS, Chon E, Park J, et al. Factors to improve distress and fatigue in Cancer survivorship; further understanding through text analysis of interviews by machine learning. BMC Cancer 2021 Jun 27;21(1):741 [FREE Full text] [doi: 10.1186/s12885-021-08438-8] [Medline: 34176470]
26. NetMiner semantic network analysis. Cyram Inc, Seongnam. 2018. URL: http://www.netminer.com/aboutus/publications.do#manual [accessed 2022-01-12]
27. Jo W, Lee J, Park J, Kim Y. Online information exchange and anxiety spread in the early stage of the novel coronavirus (COVID-19) outbreak in South Korea: structural topic model and network analysis. J Med Internet Res 2020 Jun 02;22(6):e19455 [FREE Full text] [doi: 10.2196/19455] [Medline: 32463367]
28. Fortunato S, Hric D. Community detection in networks: a user guide. Phys Rep 2016 Nov;659:1-44 [FREE Full text] [doi: 10.1016/j.physrep.2016.09.002]
29. Marjerrison S, Hendershot E, Empringham B, Nathan PC. Smoking, binge drinking, and drug use among childhood cancer survivors: a meta-analysis. Pediatr Blood Cancer 2016 Jul;63(7):1254-1263. [doi: 10.1002/pbc.25943] [Medline: 26999299]
30. Kilsdonk E, Wendel E, van Dulmen-den Broeder E, van Leeuwen FE, van den Berg MH, Jaspers MW. Participation rates of childhood cancer survivors to self-administered questionnaires: a systematic review. Eur J Cancer Care (Engl) 2017 Nov;26(6). [doi: 10.1111/ecc.12462] [Medline: 26880318]
31. Galán S, de la Vega R, Miró J. Needs of adolescents and young adults after cancer treatment: a systematic review. Eur J Cancer Care (Engl) 2018 Nov;27(6):e12558. [doi: 10.1111/ecc.12558] [Medline: 27507260]
32. Ismail Y, Hendry J. Support needs of adolescents' post-cancer treatment: a systematic review. Radiography (Lond) 2018 May;24(2):175-183. [doi: 10.1016/j.radi.2017.12.004] [Medline: 29605116]
33. Janin MM, Ellis SJ, Wakefield CE, Fardell JE. Talking about cancer among adolescent and young adult cancer patients and survivors: a systematic review. J Adolesc Young Adult Oncol 2018 Oct;7(5):515-524. [doi: 10.1089/jayao.2017.0131] [Medline: 29851370]
34. Nicklin E, Velikova G, Hulme C, Lopez R, Glaser A, Kwok-Williams M, et al. Long-term issues and supportive care needs of adolescent and young adult childhood brain tumour survivors and their caregivers: a systematic review. Psychooncology 2019 Mar;28(3):477-487. [doi: 10.1002/pon.4989] [Medline: 30657618]
35. Stone DS, Ganz PA, Pavlish C, Robbins WA. Young adult cancer survivors and work: a systematic review. J Cancer Surviv 2017 Dec;11(6):765-781 [FREE Full text] [doi: 10.1007/s11764-017-0614-3] [Medline: 28478587]
36. Ke Y, Ng T, Chan A. Survivorship care models for breast cancer, colorectal cancer, and adolescent and young adult (AYA) cancer survivors: a systematic review. Support Care Cancer 2018 Jul;26(7):2125-2141. [doi: 10.1007/s00520-018-4197-y] [Medline: 29696424]
37. Girardi F, Allemani C, Coleman MP. Worldwide trends in survival from common childhood brain tumors: a systematic review. J Glob Oncol 2019 Oct;5:1-25 [FREE Full text] [doi: 10.1200/JGO.19.00140] [Medline: 31682549]
38. Zabih V, Kahane A, O'Neill NE, Ivers N, Nathan PC. Interventions to improve adherence to surveillance guidelines in survivors of childhood cancer: a systematic review. J Cancer Surviv 2019 Oct;13(5):713-729. [doi: 10.1007/s11764-019-00790-w] [Medline: 31338733]
39. Pugh G, Gravestock HL, Hough RE, King WM, Wardle J, Fisher A. Health behavior change interventions for teenage and young adult cancer survivors: a systematic review. J Adolesc Young Adult Oncol 2016 Jun;5(2):91-105. [doi: 10.1089/jayao.2015.0042] [Medline: 26871647]

40. Mizrahi D, Wakefield CE, Fardell JE, Quinn VF, Lim Q, Clifford BK, et al. Distance-delivered physical activity interventions for childhood cancer survivors: a systematic review and meta-analysis. Crit Rev Oncol Hematol 2017 Oct;118:27-41. [doi: 10.1016/j.critrevonc.2017.08.008] [Medline: 28917267]
41. Ladhani S, Empringham B, Wang K, Portwine C, Banfield L, de Souza RJ, et al. Overweight and obesity management strategies in survivors of paediatric acute lymphoblastic leukaemia: a systematic review protocol. BMJ Open 2018 Jun 22;8(6):e022530 [FREE Full text] [doi: 10.1136/bmjopen-2018-022530] [Medline: 29934396]
42. Shliakhtsitsava K, Romero SA, Dewald SR, Su HI. Pregnancy and child health outcomes in pediatric and young adult leukemia and lymphoma survivors: a systematic review. Leuk Lymphoma 2018 Feb;59(2):381-397 [FREE Full text] [doi: 10.1080/10428194.2017.1352097] [Medline: 28728448]
43. Wang K, Fleming A, Johnston DL, Zelcer SM, Rassekh SR, Ladhani S, et al. Overweight, obesity and adiposity in survivors of childhood brain tumours: a systematic review and meta-analysis. Clin Obes 2018 Feb;8(1):55-67. [doi: 10.1111/cob.12224] [Medline: 29024558]
44. Antwi GO, Jayawardene W, Lohrmann DK, Mueller EL. Physical activity and fitness among pediatric cancer survivors: a meta-analysis of observational studies. Support Care Cancer 2019 Sep;27(9):3183-3194. [doi: 10.1007/s00520-019-04788-z] [Medline: 30993451]
45. Morales JS, Valenzuela PL, Rincón-Castanedo C, Santos-Lozano A, Fiuza-Luces C, Lucia A. Is health status impaired in childhood cancer survivors? A systematic review and meta-analysis. Crit Rev Oncol Hematol 2019 Oct;142:94-118. [doi: 10.1016/j.critrevonc.2019.07.008] [Medline: 31394434]
46. Poon LH, Yu C, Peng L, Ewig CL, Zhang H, Li C, et al. Clinical ascertainment of health outcomes in Asian survivors of childhood cancer: a systematic review. J Cancer Surviv 2019 Jun;13(3):374-396 [FREE Full text] [doi: 10.1007/s11764-019-00759-9] [Medline: 31055708]
47. Zhi X, Xie M, Zeng Y, Liu J, Cheng AS. Effects of exercise intervention on quality of life in adolescent and young adult cancer patients and survivors: a meta-analysis. Integr Cancer Ther 2019;18:1534735419895590 [FREE Full text] [doi: 10.1177/1534735419895590] [Medline: 31845599]
48. Morales JS, Valenzuela PL, Herrera-Olivares AM, Baño-Rodrigo A, Castillo-García A, Rincón-Castanedo C, et al. Exercise interventions and cardiovascular health in childhood cancer: a meta-analysis. Int J Sports Med 2020 Mar;41(3):141-153. [doi: 10.1055/a-1073-8104] [Medline: 31935777]
49. Ramsey WA, Heidelberg RE, Gilbert AM, Heneghan MB, Badawy SM, Alberts NM. eHealth and mHealth interventions in pediatric cancer: a systematic review of interventions across the cancer continuum. Psychooncology 2020 Jan;29(1):17-37. [doi: 10.1002/pon.5280] [Medline: 31692183]
50. Skiba MB, McElfresh JJ, Howe CL, Crane TE, Kopp LM, Jacobs ET, et al. Dietary interventions for adult survivors of adolescent and young adult cancers: a systematic review and narrative synthesis. J Adolesc Young Adult Oncol 2020 Jun;9(3):315-327. [doi: 10.1089/jayao.2019.0105] [Medline: 31904297]
51. Chen P, Melissa HM, Li M, Huang I. Health utilities in pediatric cancer patients and survivors: a systematic review and meta-analysis for clinical implementation. Qual Life Res 2021 Jul 05. [doi: 10.1007/s11136-021-02931-0] [Medline: 34224073]
52. Wartenberg L, Raber M, Chandra J. Unique features of a web-based nutrition website for childhood cancer populations: descriptive study. J Med Internet Res 2021 Sep 13;23(9):e24515 [FREE Full text] [doi: 10.2196/24515] [Medline: 34515643]
53. Schmidt R, Richter D, Sender A, Geue K. Motivations for having children after cancer--a systematic review of the literature. Eur J Cancer Care (Engl) 2016 Jan;25(1):6-17. [doi: 10.1111/ecc.12276] [Medline: 25545015]
54. Turner JK, Hutchinson A, Wilson C. Correlates of post-traumatic growth following childhood and adolescent cancer: a systematic review and meta-analysis. Psychooncology 2018 Apr;27(4):1100-1109. [doi: 10.1002/pon.4577] [Medline: 29096418]
55. Yang Y, Li W, Wen Y, Wang H, Sun H, Liang W, et al. Fear of cancer recurrence in adolescent and young adult cancer survivors: a systematic review of the literature. Psychooncology 2019 Apr;28(4):675-686. [doi: 10.1002/pon.5013] [Medline: 30703261]
56. De R, Zabih V, Kurdyak P, Sutradhar R, Nathan PC, McBride ML, et al. Psychiatric disorders in adolescent and young adult-onset cancer survivors: a systematic review and meta-analysis. J Adolesc Young Adult Oncol 2020 Feb;9(1):12-22. [doi: 10.1089/jayao.2019.0097] [Medline: 31674879]
57. Di Giuseppe G, Thacker N, Schechter T, Pole JD. Anxiety, depression, and mental health-related quality of life in survivors of pediatric allogeneic hematopoietic stem cell transplantation: a systematic review. Bone Marrow Transplant 2020 Jul;55(7):1240-1254. [doi: 10.1038/s41409-020-0782-z] [Medline: 31980743]
58. Sharkey CM, Espeleta HC, Traino KA, Roberts CM, Perez MN, Bakula DM, et al. Psychological adjustment outcomes among pediatric brain tumor survivors: a meta-analysis. Pediatr Blood Cancer 2020 Oct;67(10):e28644. [doi: 10.1002/pbc.28644] [Medline: 32761992]
59. Smith LA, Critoph DJ, Hatcher HM. How can health care professionals communicate effectively with adolescent and young adults who have completed cancer treatment? A systematic review. J Adolesc Young Adult Oncol 2020 Jun;9(3):328-340. [doi: 10.1089/jayao.2019.0133] [Medline: 31934804]

60. Warner EL, Kent EE, Trevino KM, Parsons HM, Zebrack BJ, Kirchhoff AC. Social well-being among adolescents and young adults with cancer: a systematic review. Cancer 2016 Apr 01;122(7):1029-1037 [FREE Full text] [doi: 10.1002/cncr.29866] [Medline: 26848713]
61. Schulte F, Russell KB, Cullen P, Embry L, Fay-McClymont T, Johnston D, et al. Systematic review and meta-analysis of health-related quality of life in pediatric CNS tumor survivors. Pediatr Blood Cancer 2017 Aug;64(8). [doi: 10.1002/pbc.26442] [Medline: 28266804]
62. Collins DE, Ellis SJ, Janin MM, Wakefield CE, Bussey K, Cohn RJ, et al. A Systematic Review Summarizing the State of Evidence on Bullying in Childhood Cancer Patients/Survivors [Formula: see text]. J Pediatr Oncol Nurs 2019;36(1):55-68. [doi: 10.1177/1043454218810136] [Medline: 30406714]
63. Molcho M, D'Eath M, Alforque Thomas A, Sharp L. Educational attainment of childhood cancer survivors: A systematic review. Cancer Med 2019 Jun;8(6):3182-3195 [FREE Full text] [doi: 10.1002/cam4.2154] [Medline: 31007000]
64. Kelada L, Wakefield CE, Silva MC, Signorelli C. Camps for children with cancer and their families: a systematic review of psychosocial and physical impacts. J Dev Behav Pediatr 2020;41(2):145-156. [doi: 10.1097/DBP.0000000000000728] [Medline: 31567721]
65. Saatci D, Thomas A, Botting B, Sutcliffe AG. Educational attainment in childhood cancer survivors: a meta-analysis. Arch Dis Child 2020 Apr;105(4):339-346. [doi: 10.1136/archdischild-2019-317594] [Medline: 31601570]
66. Gerstl B, Sullivan E, Chong S, Chia D, Wand H, Anazodo A. Reproductive Outcomes After a Childhood and Adolescent Young Adult Cancer Diagnosis in Female Cancer Survivors: A Systematic Review and Meta-analysis. J Adolesc Young Adult Oncol 2018 Nov 16. [doi: 10.1089/jayao.2018.0036] [Medline: 30452301]
67. Tennyson RE, Griffiths HC. A systematic review of professionals' experiences of discussing fertility issues with adolescents and young adults with cancer. J Adolesc Young Adult Oncol 2019 Aug;8(4):387-397. [doi: 10.1089/jayao.2018.0146] [Medline: 30900925]
68. Bhurosy T, Niu Z, Heckman CJ. Breastfeeding is possible: a systematic review on the feasibility and challenges of breastfeeding among breast cancer survivors of reproductive age. Ann Surg Oncol 2021 Jul;28(7):3723-3735. [doi: 10.1245/s10434-020-09094-1] [Medline: 32915334]
69. Jibb LA, Croal L, Wang J, Yuan C, Foster J, Cheung V, et al. Children's Experiences of Cancer Care: A Systematic Review and Thematic Synthesis of Qualitative Studies. Oncol Nurs Forum 2018 Jul 02;45(4):527-544. [doi: 10.1188/18.ONF.527-544] [Medline: 29947357]
70. Teepen JC, de Vroom SL, van Leeuwen FE, Tissing WJ, Kremer LC, Ronckers CM. Risk of subsequent gastrointestinal cancer among childhood cancer survivors: A systematic review. Cancer Treat Rev 2016 Feb;43:92-103. [doi: 10.1016/j.ctrv.2015.12.002] [Medline: 26827697]
71. Kurt I, Tanriverdi M. New approaches of oncology rehabilitation in children: technotherapy-systematic literature review. Neuro-Oncology 2018;20(Suppl 2):i166. [doi: 10.1093/neuonc/noy059.626]
72. Benetou D, Stergianos E, Geropeppa M, Ntinopoulou E, Tzanni M, Pourtsidis A, et al. Late-onset cardiomyopathy among survivors of childhood lymphoma treated with anthracyclines: a systematic review. Hellenic J Cardiol 2019;60(3):152-164 [FREE Full text] [doi: 10.1016/j.hjc.2018.09.004] [Medline: 30273645]
73. Yildiz Kabak V, Calders P, Duger T, Mohammed J, van Breda E. Short and long-term impairments of cardiopulmonary fitness level in previous childhood cancer cases: a systematic review. Support Care Cancer 2019 Jan;27(1):69-86. [doi: 10.1007/s00520-018-4483-8] [Medline: 30251066]
74. Leerink JM, Verkleij SJ, Feijen EA, Mavinkurve-Groothuis AM, Pourier MS, Ylänen K, et al. Biomarkers to diagnose ventricular dysfunction in childhood cancer survivors: a systematic review. Heart 2019 Feb;105(3):210-216. [doi: 10.1136/heartjnl-2018-313634] [Medline: 30158136]
75. Meijer AJ, Clemens E, Hoetink AE, van Grotel M, van den Heuvel-Eibrink MM. Tinnitus during and after childhood cancer: a systematic review. Crit Rev Oncol Hematol 2019 Mar;135:1-7. [doi: 10.1016/j.critrevonc.2019.01.004] [Medline: 30819438]
76. Pelland-Marcotte M, Hwee J, Pole JD, Nathan PC, Sung L. Incidence of infections after therapy completion in children with acute lymphoblastic leukemia or acute myeloid leukemia: a systematic review of the literature. Leuk Lymphoma 2019 Sep;60(9):2104-2114. [doi: 10.1080/10428194.2019.1573369] [Medline: 30774019]
77. Rustler V, Däggelmann J, Streckmann F, Bloch W, Baumann FT. Whole-body vibration in children with disabilities demonstrates therapeutic potentials for pediatric cancer populations: a systematic review. Support Care Cancer 2019 Feb;27(2):395-406. [doi: 10.1007/s00520-018-4506-5] [Medline: 30368670]
78. Paetow U, Bader P, Chemaitilly W. A systematic approach to the endocrine care of survivors of pediatric hematopoietic stem cell transplantation. Cancer Metastasis Rev 2020 Mar;39(1):69-78. [doi: 10.1007/s10555-020-09864-z] [Medline: 31980968]
79. Peng L, Yam PP, Yang LS, Sato S, Li CK, Cheung YT. Neurocognitive impairment in Asian childhood cancer survivors: a systematic review. Cancer Metastasis Rev 2020 Mar;39(1):27-41. [doi: 10.1007/s10555-020-09857-y] [Medline: 31965433]

Abbreviations

LDA: latent Dirichlet allocation
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QOL: quality of life
TF-IDF: term frequency-inverse document frequency

Edited by A Mavragani; submitted 22.07.21; peer-reviewed by B Zhang, C Walsh; comments to author 17.08.21; revised version received 06.10.21; accepted 08.12.21; published 01.02.22.

Please cite as:
Kim HY, Kang KA, Han SJ, Chun J
Web-Based Research Trends on Child and Adolescent Cancer Survivors Over the Last 5 Years: Text Network Analysis and Topic Modeling Study
J Med Internet Res 2022;24(2):e32309
URL: https://www.jmir.org/2022/2/e32309
doi:10.2196/32309
PMID:35103615

Original Paper

Longitudinal Changes of COVID-19 Symptoms in Social Media: Observational Study

Sarah Sarabadani[1], MASc; Gaurav Baruah[1], DPhil; Yan Fossat[1], BSc; Jouhyun Jeon[1], DPhil

Applied Sciences, Klick Inc, Toronto, ON, Canada

Corresponding Author:
Jouhyun Jeon, DPhil
Applied Sciences
Klick Inc
175 Bloor Street East, Suite 300
Toronto, ON, M4W 3R8
Canada
Phone: 1 4162144977
Email: cjeon@klick.com

Abstract

Background: In December 2019, the COVID-19 outbreak started in China and rapidly spread around the world. Many studies have been conducted to understand the clinical characteristics of COVID-19, and recently postinfection sequelae of this disease have begun to be investigated. However, there is little consensus on the longitudinal changes of lasting physical or psychological symptoms from prior COVID-19 infection.

Objective: This study aims to investigate and analyze public social media data from Reddit to understand the longitudinal impact of COVID-19 symptoms before and after recovery from COVID-19.

Methods: We collected 22,890 Reddit posts that were generated by 14,401 authors from March 14 to December 16, 2020. Using active learning and intensive manual inspection, 292 (2.03%) active authors, who were infected by COVID-19 and frequently reported disease progress on Reddit, along with their 2213 (9.67%) longitudinal posts, were identified. Machine learning tools to extract biomedical information were applied to identify COVID-19 symptoms mentioned in the Reddit posts. We then examined longitudinal changes in individual physiological and psychological characteristics before and after recovery from COVID-19 infection.

Results: In total, 58 physiological and 3 psychological symptoms were identified in social media before and after recovery from COVID-19 infection. From the analyses, we found that symptoms of patients with COVID-19 lasted 2.5 months. On average, symptoms appeared around a month before recovery and remained for 1.5 months after recovery. Well-known COVID-19 symptoms, such as fever, cough, and chest congestion, appeared relatively earlier in patient journeys and were frequently observed before recovery from COVID-19. Meanwhile, mental discomfort or distress, such as brain fog or stress, fatigue, and manifestations on toes or fingers, were frequently mentioned after recovery and remained as intermediate- and longer-term sequelae.

Conclusions: In this study, we showed the dynamic changes in COVID-19 symptoms during the infection and recovery phases of the disease. Our findings suggest the feasibility of using social media data for investigating disease states and understanding the evolution of the physiological and psychological characteristics of COVID-19 infection over time.

(J Med Internet Res 2022;24(2):e33959) doi:10.2196/33959

KEYWORDS

COVID-19; symptom; diagnosis; treatment; social media; Reddit; longitudinal; observational; machine learning

Introduction

COVID-19 is a pandemic viral infectious disease that has quickly spread worldwide. The clinical presentation of COVID-19 has been well defined from active clinical and basic scientific research. Risk factors and commonly observed symptoms at the diagnosis of COVID-19 and throughout the acute disease course have been well documented [1,2]. Several studies have tried to identify self-reported COVID-19 symptoms using social media data [3,4] and investigated the dynamics of symptoms that were observed prior to and throughout COVID-19 infection [5]. Our own analysis of social media and clinical literatures suggested less common or rarely observed

novel symptoms related to COVID-19. We also observed that different sets of clinical and demographic characteristics are associated with specific clinical outcomes, such as severity of disease progression, hospitalization, and intensive care unit admission [6]. These efforts have helped to understand the development of COVID-19 and identify appropriate medical treatment options and diagnostic methods.

Since early 2021, researchers have found emerging evidence of long-term sequelae in a considerable proportion of patients who have recovered from COVID-19. Several systematic reviews and cohort-based studies have suggested a set of symptoms from which COVID-19 survivors have partially recovered or those they have retained [7,8]. Most of these studies have focused on a set of symptoms that appeared at a static time point (eg, symptoms before, or at, the moment of confirmation of COVID-19; symptoms during recovery; or symptoms after recovery from COVID-19) and were based on retrospective data of hospitalized patients. When we considered that about 15% of patients with COVID-19 have been hospitalized [9], there was limited understanding of patients who experienced less severe symptoms and received home-based care. This created a huge knowledge gap to understand the symptom landscape and symptom durations in the general patient population (ie, 85% of patients with COVID-19). A comprehensive understanding of the full spectrum of symptoms and their dynamic changes throughout the patient journey (ie, disease course) was essential to obtain better information about acute and chronic/persistent COVID-19 symptoms in general patients and develop a complete understanding of the longitudinal impact of COVID-19 infection in the population.

To better understand the longitudinal changes in the physiological and psychological characteristics of the COVID-19 patient population throughout the patient journey, we systematically investigated Reddit public social media data from individuals who had previously been infected with COVID-19 and have recovered. Social media provides an efficient method of gathering large amounts of real-world data on the general public, which are scalable and convenient for users at any time of day, especially from remote or unattended regions. In addition, the development of sophisticated machine learning methods has been used to collect high-quality medical information from social media [10,11]. For this study, social media data were utilized to capture the disease course of general patients with COVID-19, including nonhospitalized patients with various levels of symptom severity. Machine learning methods were applied to select reliable patients with COVID-19 and their posts from social media data and to identify the full spectrum of symptoms of COVID-19 in a comprehensive manner.

Methods

Data

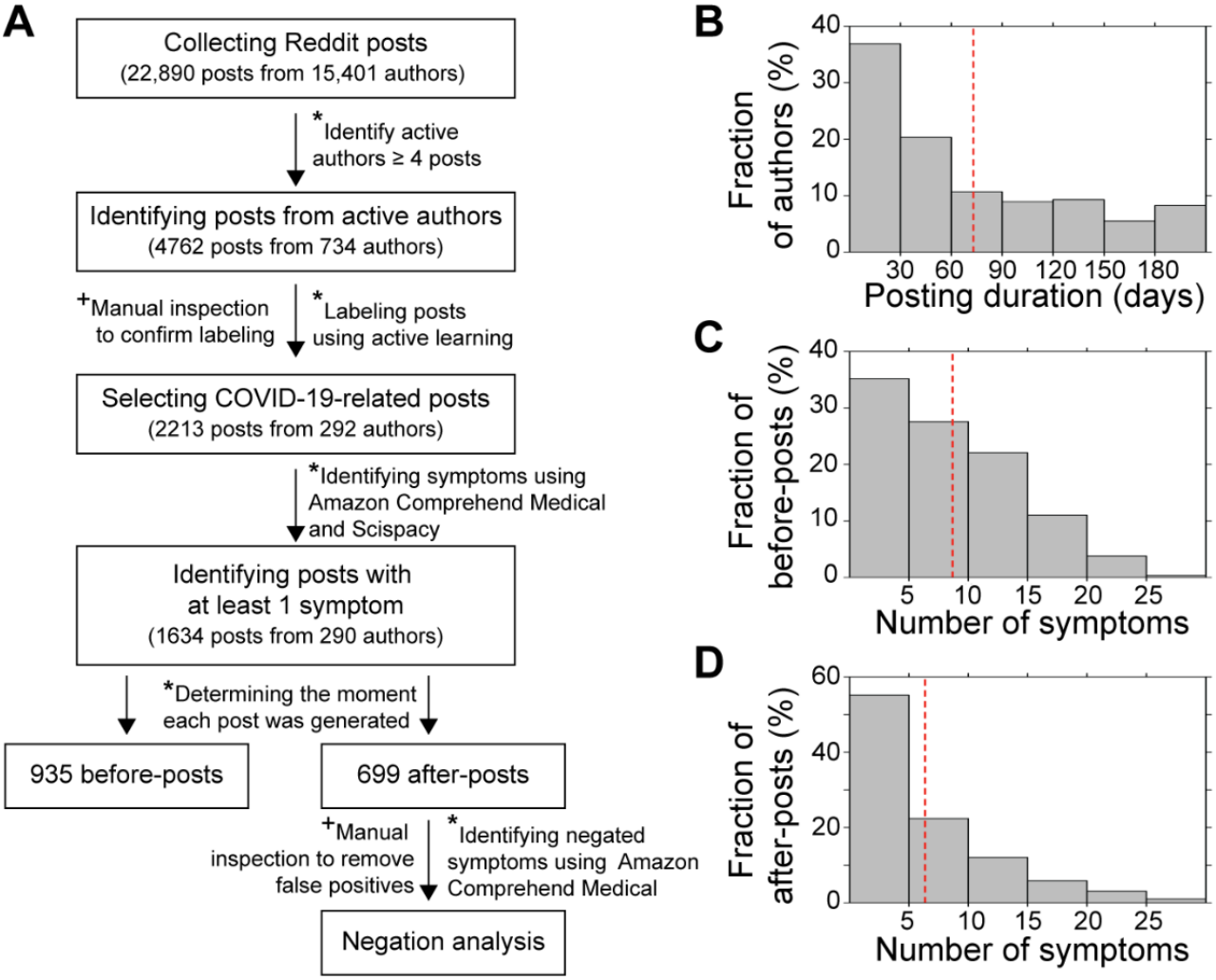
Reddit was used to collect social media posts of COVID-19 survivors. Reddit is a large user-generated content website and text-based social media platform (without the limitation of text length) enabling us to extract detailed information about a given topic. Reddit is also a community-based social media platform providing sections devoted to topic-specific discussions (subreddits). Authors express their opinions or concerns on subreddit forums. Subreddit defines its characteristics and eligible members. Therefore, people who are really interested in each topic can participate in a subreddit community. Furthermore, Reddit provides a publicly accessible Application Programming Interface (API) so that researchers can collect and analyze anonymized posts for their own purpose. Such clear definitions of topics and member eligibility, as well as anonymized data access through the API, made Reddit a reliable and effective social media platform to find the appropriate target audience for this study.

For this analysis, we collected posts that were generated from the moment the United States declared COVID-19 a national emergency (March 13, 2020; we collected posts from March 14, 2020) until the Food and Drug Administration (FDA) issued an emergency use authorization to both Pfizer-BioNTech's and Moderna's COVID-19 vaccines (December 17, 2020; we collected posts until December 16, 2020). After the World Health Organization (WHO) declared COVID-19 a pandemic and the United States declared COVID-19 a national emergency, it was logical to assume that the general population started to become aware of the presence of the COVID-19 virus and be interested in symptoms and their health conditions. In addition, they were exposed to the risk of COVID-19 contraction before the vaccine was available. During this period, people could actively discuss COVID-19-related issues and share their personal experience. We especially considered posts generated in "COVID-19Positive" and "CoronavirusSurvivors" subreddits. "COVID-19Positive" was self-described as "a place for people who came back positive for COVID-19 to share your stories, experiences, answer questions and vent!" "CoronavirusSurvivor" was self-described as a community for survivors of the coronavirus. At the time the conversations analyzed for this study were downloaded, the "COVID-19Positive" subreddit had about 81,000 members and the "CoronavirusSurvivors" subreddit had about 1000 members.

The procedure to process and analyze social media data is described in Figure 1A. In total, 22,890 posts from 15,401 authors were collected from Reddit. Of the 15,401 authors, 11,900 (77.27%) generated only 1 post, which was inappropriate to examine longitudinal changes in COVID-19 symptoms. We did not consider them for further analyses.

To identify active authors' authentic COVID-19-related posts, we performed active learning. After the automated active learning procedure, manual inspection was performed to examine the reliability of post labeling (Figure 1A). Details are described in the Acquiring Pertinent Labels for Reddit Posts section.

Figure 1. Overview of social media data related to COVID-19 patient journey. (A) Social media data collection and analysis procedure. In total, 1634 posts from 290 active authors who continuously posted their disease progress were identified and used for further analyses. * and + indicate automated steps and steps requiring manual inspections, respectively. (B) Fraction of active authors, depending on posting duration. Posting duration is the time difference between the first and last COVID-19 posts from the same active author. The red-dashed line indicates the average posting duration. (C) Fraction of before-posts, depending on number of symptoms. (D) Fraction of after-posts, depending on the number of symptoms. The red-dashed line indicates the average number of symptoms mentioned in a post.



Acquiring Pertinent Labels for Reddit Posts

We observed some authors asking questions about COVID-19 or posting about symptoms, or about a family member, rather than themselves. It was apparent after reading many posts that not all authors were COVID-19 positive. This observation necessitated distinguishing between true patients with COVID-19 and authors who had not tested positive for COVID-19. Furthermore, we wanted to identify those authors that had discussed recovery from COVID-19 in their posts. We were also interested to know how different physical and psychological symptoms were discussed within Reddit posts.

We first generated an unlabeled data set of 4762 posts that were written by 734 authors who had posted at least 4 times in our COVID-19-related subreddits (Figure 1A). We designated 4 classes (ie, label types) that we needed for our analysis. For each post in our data set, we needed to check whether the post had (1) evidence that the author is/was COVID-19 positive (label: “posi”), (2) evidence that the author has recovered from COVID-19 (label: “reco”), (3) evidence of physiological symptoms (label: “physio”), and (4) evidence of psychological symptoms (label: “psycho”). More details of labeling are described in Multimedia Appendix 1.

To do this, we implemented an active learning method [12] to capture the different labels for Reddit posts in an automated and expedited manner. Active learning is a popular method to find relevant materials from within documents and is widely used in the e-discovery domain (eg, labeling documents as relevant or nonrelevant from a collection of legal documents or research papers). It also has been applied to identify the literature that is relevant to infection prevention and control of COVID-19 from a large biomedical corpus [13] and used to build a diagnostic method for COVID-19 [14].

Using active learning, we built a classifier for each of the label classes; the classifier was periodically updated with new evidence, as saved by a human assessor (Multimedia Appendix 1). This way, the classifier can capture newly found text features with each training iteration. The goal for the classifier is to find the next best-candidate Reddit post that can contain a label evidence string. The human assessor, when presented with the candidate post, assesses the post as nonrelevant (providing a

true-negative label) or highlights and saves an evidence string (providing a true-positive label). After retraining with a set of new labels, the classifier "relearns" which features in the text are likely to be found relevant and it presents the next best candidate to the user for assessment. For the human assessor, the likelihood of finding true-positive labels increases early in the label-gathering process, which helps with expediting labeling efforts when we are looking for true-positive labels on a budget.

A labeling interface was built to capture labels using our active learning method (Multimedia Appendix 1). An assessor is presented with 1 Reddit post at a time for 1 label class. The assessor can mark the post as nonrelevant for the particular post for the particular class or save (copy/paste) an evidence string that is representative of the label class description. The classifier for the respective label type is then retrained with the new evidence string. Then, all remaining unlabeled posts are scored by the classifier. The top-ranked unlabeled post is then presented to the assessor for judgement. Thus, after each judgement cycle, the classifier learns which text features are relevant to the assessor, and the method then presents the most likely relevant candidate for assessment to the user next. This way, the most relevant posts are labeled quickly, with the added advantage that assessment budgets and available time are utilized more efficiently.

We recruited 13 well-trained assessors for the labeling task; the number of assessors varied across label classes. The goal for assessors was to find and label as many label-relevant posts as fast as possible. To do this, we provided 2 hours of training sessions so that accessors could find relevant posts generated by reliable patients with COVID-19. Labeled data completed by accessors were then manually inspected by the authors in this study to confirm that labeling was done correctly (Multimedia Appendix 1). The assessors made 2785 judgements and found 1072 evidence strings across all classes (Multimedia Appendix 1).

Identifying Active Authors

We selected active authors who were infected by COVID-19 and described their experiences and COVID-19 symptoms across the entire patient journey: from the confirmation of COVID-19 infection to after recovery from COVID-19. To do this, we examined the median time duration between the first post and the last post that a given user generated (Multimedia Appendix 1), and selected the optimal number of posts that could represent the patient journey.

We found that symptoms usually take 5-6 days to appear after COVID-19 exposure (incubation period of COVID-19) [15]. We also estimated the recovery time after onset of the symptoms. We calculated the average recovery time (duration between the posts with the label "posi" and the posts with the label "reco") for the subset of users who had mention of recovery in their posts. For this calculation, first we removed the outliers that were considered as cases with recovery time higher than 2×SD. We obtained 15.7(SD 10.4) days as the average recovery time after testing positive. Since our subset was small (n=38 users) and there was high variation in the data, we decided to use the commonly accepted recovery time from the literature. Although there were various studies reporting different durations [5], 14 days was the duration we most often found in our search [16-19]. Therefore, we utilized 14 days as the recovery period after having tested positive for COVID-19, and we assumed that 20 (6+14) days of posting period could represent a typical patient's journey. In addition, 4 posts were used as a cut-off to find active authors since they were generated for 22 days, which was the closest period to our adopted notion of COVID-19 duration (20 days).

Identifying Symptoms of COVID-19

To identify the physical or psychological symptoms that were mentioned in Reddit posts, we applied 2 automated symptom extraction methods. Using the Amazon Comprehend Medical tool (Amazon Web Services), we extracted medical entities. The Amazon Comprehend Medical tool uses machine learning to extract health-related information from text automatically. For this study, we considered medical entities "symptoms" and "signs" as COVID-19 symptoms. In addition to the Amazon Comprehend Medical tool, we also built a model to identify medical entities using Scispacy (v.0.4.0). Scispacy is a Python package for handling scientific documents and extracting medical and clinical terminology [20]. We considered the medical entity "disease" as COVID-19 symptoms. From the performance evaluation of medical entity extraction models, we found that the models identified over 80% of COVID-19 symptoms correctly and reliably (Multimedia Appendix 1). The model achieved 87% precision, 83% recall, and an F1 score of 0.85. In total, 58 physical and 3 psychological symptoms were identified (Multimedia Appendix 1). The 3 psychological symptoms were "confusion or fluster," "depression or anxiety," and "mental discomfort or distress" (eg, foggy head and loss of consciousness).

In total, 1634 posts generated by 290 active authors mentioned at least 1 COVID-19 symptom (Figure 1A). Next, we determined when each post was generated: before or after recovery from COVID-19. We divided the posts into 2 groups, before-posts and after-posts, based on their posting time. Before-posts were written before COVID-19 recovery, and after-posts were written after recovery. Posts that were generated before recovery posts (label: "reco") were defined as before-posts. When authors only had positive posts (label: "posi"), posts generated from the date of the first positive post to the next 14 days were considered as before-posts based on the study that patients with COVID-19 take about 14 days to recover from the disease [16]. Remaining posts were defined as after-posts. We assumed the date of the first before-post was the moment users realized COVID-19 symptoms or confirmed their infection by COVID-19.

Since the number of before- and after-posts was different, identified COVID-19 symptoms were observed with different frequencies. For a fair comparison of symptom frequency between before and after COVID-19 recovery, we normalized the frequency of each symptom to 100 posts (ie, of 100 posts, how many mentioned a given symptom). Depending on this frequency, acute and chronic symptoms were determined. Acute symptoms developed rapidly and were mentioned more frequently in before-posts. Chronic symptoms developed gradually and were slow to resolve, remaining as sequelae, and

were mentioned more frequently after recovery from COVID-19 (after-posts).

Identifying Symptoms Mentioned Together

To identify symptoms that commonly appear together, we performed association rule analysis [21]. We measured support, which is defined as the proportion of posts in which a certain set of symptoms come together. We adopted the Apriori algorithm and followed a bottom-up approach that starts from every single symptom, and then symptom subsets are extended 1 item at a time. At each step, the group of candidates is tested, and the ones that include infrequent items are pruned.

Negation Analysis

To understand how patients perceive and respond to COVID-19 symptoms, negation analysis was performed using the Amazon Comprehend Medical negation model [22] and manual inspection (Figure 1A). Negation analysis showed whether a given symptom was denied (eg, "I have gotten no fever" or "I don't cough anymore"). Therefore, in our study, negation indicated that a given symptom was relieved or disappeared after recovery from COVID-19. After automatic identification of negated symptoms, we removed false positives through manual inspection. False positives are symptoms that present after COVID-19 recovery, but negation analysis detected them due to the sentence structure (eg, "I still have no smell and taste after I received a negative polymerase chain reaction [PCR] test"; "no smell and taste" was not negation. This symptom was still presented after recovery). Of 699 after-posts, negation was detected from 240 (34.3%) posts.

Statistical Analysis

To compare the duration of symptoms before and after COVID-19, Spearman rank correlation coefficient measurement was performed using the programming language Python (v.3.8.1) and SciPy package (v.1.6.2). $P<.05$ was considered statistically significant. For the visualization of analyses, the BPG library (v.6.0.1) in R was used [23].

Results

Result 1: Overview of Reddit COVID-19 Posts

To understand longitudinal changes in physical and psychological characteristics during the COVID-19 patient journey, we decided to collect patients' self-generated posts from Reddit. Reddit is a community-based online forum where patients can share their clinical journey, symptoms, and experiences from diagnosis to postrecovery. We identified 292 active authors who continuously and voluntarily discussed their physiological and psychological symptoms from the beginning of COVID-19 contraction to the after-recovery stage of COVID-19 (see the Methods section and Multimedia Appendix 1 for details). From the active learning experiment, we found that active authors generated 2213 posts that described their positive COVID-19 infection and the symptoms they had before and after recovery from COVID-19. On average, active authors generated 8 posts in 73 days (difference between the first before-post and the last after-post; Figure 1B). It is suggested that COVID-19 symptoms remain after recovery and present for about 2.5 months from diagnosis.

Next, we fed all 2213 posts into the biomedical named entity recognition (Bio-NER) program to find COVID-19 symptoms in a comprehensive manner. In total, 1634 (73.84%) posts mentioned at least 1 COVID-19 symptom (Figure 1A). From those, 935 (57.22%) posts were written from diagnosis (eg, confirmation of a positive test or notification of COVID-19 symptoms) to before recovery from COVID-19 (eg, received a negative test after the positive test). These were defined as before-posts. In addition, 699 (42.78%) posts were written after recovery (after-posts; see the Methods section for details). On average, each active author mentioned 9 symptoms in the before-post (Figure 1C) period and 6 symptoms in the after-post (Figure 1D) period.

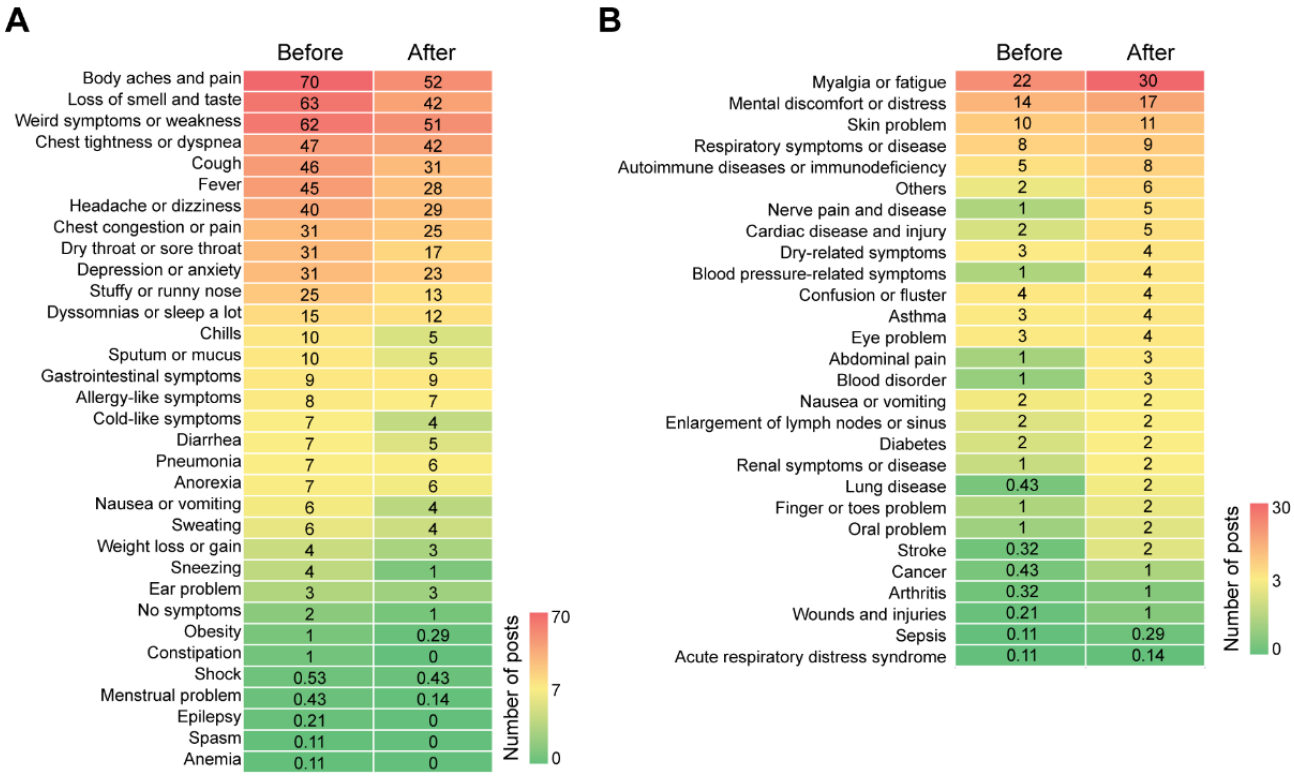
Result 2: COVID-19 Symptom Landscape in Social Media Data

To understand the dynamic development of physiological and psychological symptoms throughout the COVID-19 patient journey, we compared how often individual symptoms were mentioned in before- and after-posts. Of 61 symptoms, 33 (54%) were acute symptoms. They were mentioned more frequently before recovery from COVID-19 compared to after recovery. The most known COVID-19 symptoms, such as weakness, body aches, cough, fever, chest tightness, and loss of smell and taste, were mentioned frequently in before-posts. We observed that these symptoms were relieved or improved after recovery. For example, in 100 posts, loss of smell and taste was mentioned 63 times in before-posts. However, after recovery, it was 32.89% less frequently mentioned (42 times in after-posts). Fever, cough, and throat discomfort (dry or sore throat) were approximately 40% less frequently mentioned after recovery (Figure 2A). Chills (51.23% reduced), cold-like symptoms (42.38% reduced), nausea (40.04% reduced), and sputum (43.52% reduced) were also mentioned less after recovery. Negation analysis supported these findings. When negation was detected in a sentence that contained a symptom, we could assume that a symptom was relieved or disappeared. We performed negation analysis using after-posts and found that common symptoms, including fever, cough, or chest tightness, were denied (eg, "I have no fever or no cough") in after-posts. In addition, 5%-20% of after-posts clearly mentioned that these symptoms disappeared after recovery from COVID-19 infection (Multimedia Appendix 1). In addition, 4 (7%) extremely rare symptoms (epilepsy, spasm, constipation, and anemia; less than 1% of posts mentioned these symptoms) were only identified before the recovery period (Figure 2A).

Furthermore, 28 (46%) of 61 symptoms were mentioned more frequently after recovery from COVID-19 (chronic symptoms). We observed that patients with COVID-19 mainly complained of psychological symptoms during this time. Mental discomfort and distress, including brain fog, stress, and panic, were mentioned 18.27% more often, and confusion/fluster was mentioned 8.63% more after recovery. Constitutional symptoms, such as fatigue (25.96%), and physiological symptoms, including problems in toes or fingers (eg, tingly hands/feet and toe/finger rash, 41.26%), renal-related problems (eg, frequent

urination and kidney pain, 38.43%), and immunodeficiency (eg, autoimmune disease and immune system disorder, 37.94%), were also frequently mentioned after recovery (Figure 2B). These results suggested that well-known symptoms are likely to be acute symptoms and do not remain as postrecovery sequelae of COVID-19 infection. Rather, after recovery, patients experience mental discomfort and symptoms that are less common at or shortly following the time of diagnosis of infection and slow to resolve [6].

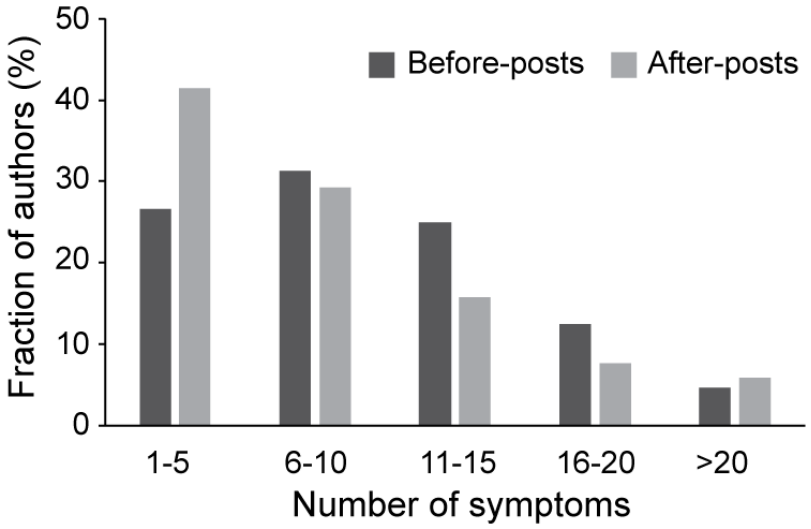
Figure 2. Symptom landscape in before- and after-posts. (A) There were 33 acute symptoms that were frequently mentioned before recovery from COVID-19. (B) There were 28 chronic symptoms that were frequently mentioned after recovery from COVID-19. The number indicates the incidence of a given symptom per 100 posts.



Result 3: Co-occurrence of COVID19 Symptoms Before and After Recovery

It has been shown that COVID-19 can cause multiple symptoms during early infection and progression [6]. To understand the dynamic co-occurrence of COVID-19 symptoms through the patient journey, we examined a set of symptoms that were mentioned together in a Reddit post. We found that each author mentioned more symptoms before recovery than after recovery (Figure 3). For example, of 256 authors who generated before-posts, 64 (25%) mentioned 11-15 symptoms, and this was 1.63 times higher compared to those who generated after-posts (n=222; 35 [15.8%] mentioned 11-15 symptoms). Meanwhile, 92 (41.4%) of 222 authors mentioned fewer symptoms (1-5 symptoms) in after-posts, which was 1.56 times higher (68/256, 26.6%) compared to before-posts. Next, we examined what symptom pairs were observed before and after recovery. In total, we identified 368 co-occurred symptom pairs. We observed that common and acute symptoms of COVID-19 were mentioned more frequently together before recovery. Chills with loss of smell and taste (2.51 times), cough (2.44 times), fever (2.39 times), chest tightness (2.39 times), or body aches (2.31 times) frequently co-occurred before recovery from COVID-19 compared to after. Meanwhile, 1 of the chronic symptoms, immunodeficiency (eg, autoimmune disease and immune system disorder), co-occurred with other chronic symptoms, such as mental discomfort/distress, myalgia/fatigue, and skin problems only after recovery from COVID-19 (Multimedia Appendix 1).

Figure 3. Number of COVID-19 symptoms in before- and after-posts. Dark- and light-gray bars indicate the fraction of authors who mentioned a given number of symptoms in before- and after-posts, respectively.



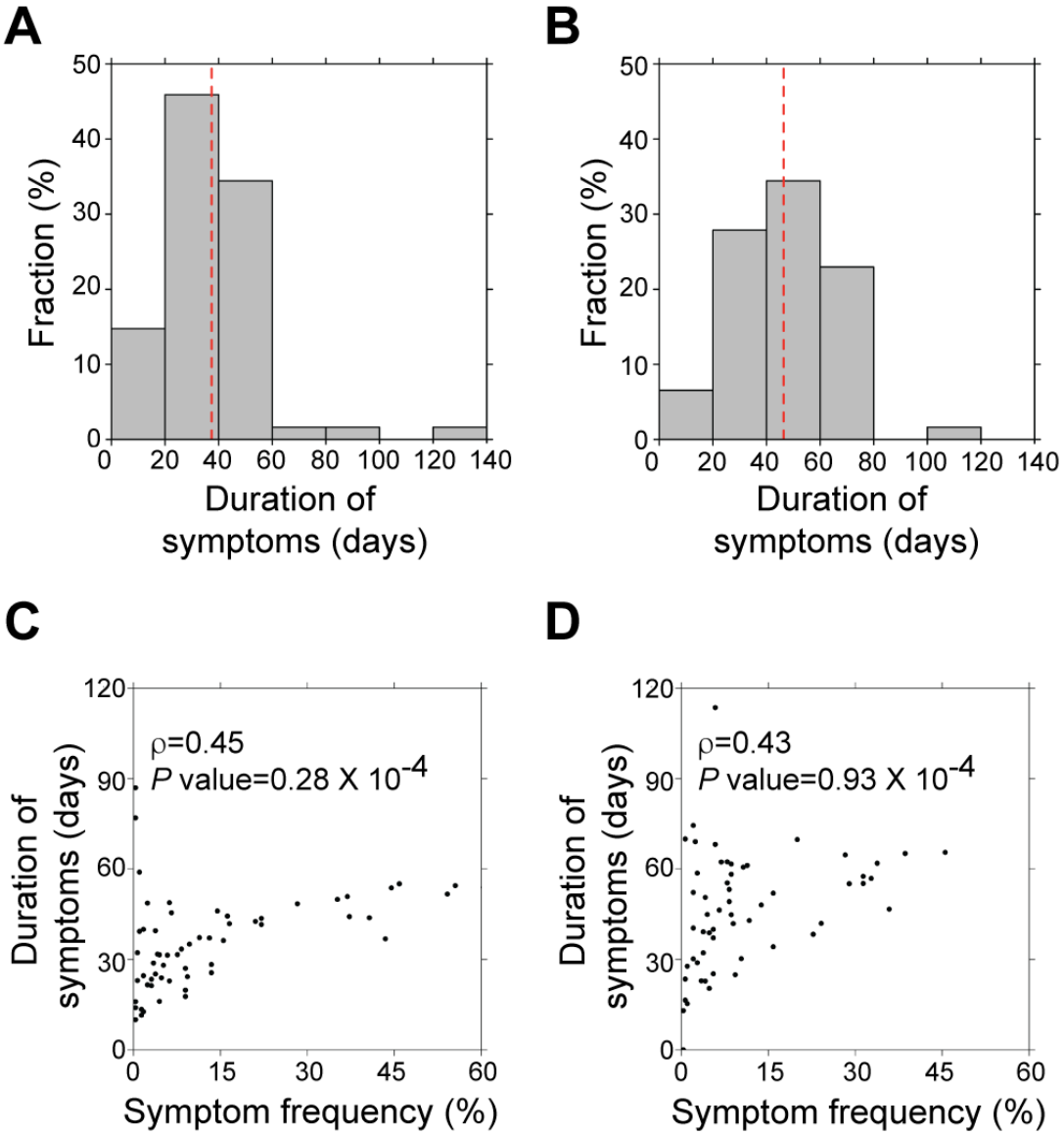
Result 4: Symptom Duration in Patients with COVID-19

Our data set was composed of posts that followed time courses, enabling us to perform temporal analyses to understand dynamic changes over time. We examined the duration of symptoms across the COVID-19 patient journey. On average, each symptom persisted for 83 days. On average, symptoms appeared 37 days before recovery (Figure 4A) and remained up to 46 days after recovery (Figure 4B). This implied that about a month was required to recover from the first COVID-19 symptom presentation and that symptoms remained for 1.5 months after recovery.

We found that there was a weak correlation between frequency of symptoms and symptom duration (ρ=0.45, P=.28×10^{-4}; Spearman correlation coefficient, Figure 4C). In before-posts, well-known acute COVID-19 symptoms (eg, weakness, body ache, cough, fever, dyspnea, and headache) were mentioned by at least 116 (40%) of 290 active authors. They appeared relatively early compared to other symptoms (appeared 54 days before recovery). Loss of smell and taste (37 days), nausea (37 days), and asthma (31 days) appeared about 1 month before recovery. Less common symptoms, including arthritis, epilepsy, and anemia (<1% of before-posts mentioned these) appeared around 10 days before recovery (Multimedia Appendix 1).

After recovery, we observed a similar trend. There was a weak but statistically significant positive correlation between frequency of symptoms and time of presentation (ρ=0.43, P=.93×10^{-4}; Spearman correlation coefficient, Figure 4D). Chronic symptoms (frequently observed after recovery), such as confusion, immunodeficiency, and fatigue, were slow to resolve. They remained for over 50 days after recovery. Mental discomfort/distress (38 days) also remained longer than other symptoms. Menstrual problems, shock, and acute respiratory distress–related symptoms (<1% of after-posts) remained about 15 days after recovery (Multimedia Appendix 1).

Figure 4. Duration of COVID-19 symptoms. (A) Duration of symptoms before recovery from COVID-19. (B) Duration of symptoms after recovery from COVID-19. The red-dashed line indicates the average symptom duration. The relationship between symptom duration and symptom occurrence (C) before and (D) after recovery from COVID-19.



Discussion

Principal Findings

Our longitudinal analysis based on social media data showed the dynamic evolution of symptoms through the COVID-19 patient journey. From this observational study, we identified acute and chronic symptoms that were frequently or specifically observed before and after recovery from COVID-19. Individual symptoms showed differing durations and recovery times. Social media data expanded our understanding of COVID-19 symptoms and their longitudinal changes through the patient journey.

From social media data, we observed that the most common COVID-19 symptoms (eg, fever, cough, and weakness) appeared earlier in patient journeys and were mentioned more frequently before recovery from COVID-19. These common COVID-19 symptoms were easily recognizable and evaluated by ordinary people without any screening tools. In addition, they were basic indicators used in clinical and medical research. Accumulated COVID-19 research defined them as well-characterized diagnostic indicators for COVID-19 [24]. Furthermore, at the beginning of the pandemic, national- or international-level awareness campaigns were conducted with a limited understanding of COVID-19 symptoms. We suggested that the public's basic awareness and easy recognition could disproportionate the prevalence of COVID-19 symptoms, increasing the reporting of common symptoms in social media before recovery from COVID-19.

Interestingly, common symptoms were relieved or disappeared after recovery from COVID-19 (acute symptoms) based on our analyses of symptom negation and the duration over this entire period. Fever was considered a beneficial response to infection. Increased temperature reduces pathogens' survival and increases mobilization of immune cells [25]. Cough is an intrinsic and protective reaction to many respiratory infections [26]. Skeletal muscle atrophy can be caused by an immune response, leading to weakness or body ache [27,28]. Taken together, we suspected that many common and acute symptoms are likely to be

associated with the initial immune response and are relieved as the initial immune response decreases over time after viral clearance.

Significance of Mining Social Media Data

Currently, there is limited information about nonhospitalized or initially asymptomatic patients with COVID-19 who have had persistent and chronic symptoms after their recovery. In addition, there is a lack of information about longitudinal changes in COVID-19 symptoms due to the limited methods or accessibility to identify COVID-19 survivors on a global level. Our study explored dynamic changes in COVID-19 symptoms throughout patient journeys using social media data. Although social media may lack some depth of patient information, it provides an effective method of collecting a wide breadth of data. Social media data can be easily gathered across the world 24 hours a day, without the need for a clinician visit, and is an extremely efficient method [29] for rapidly disseminating new knowledge related to COVID-19 [30]. Indeed, we observed more than 60 symptoms that were extracted from the Reddit posts, including all the symptoms of COVID-19 suggested by the Centers for Disease Control and Prevention (CDC) [31]. This number of symptoms observed in social media was about 2 times higher than the number of symptoms mentioned in the biomedical literature (34 symptoms), which was published by clinical institutes [6]. Furthermore, tracking COVID-19 symptoms in social media data over time gave us novel insights to understand the full clinical spectrum of symptoms and the patient journey. Social media has also been used to predict COVID-19 waves [32], forecast the number of cases [33], and develop crisis management strategies [34]. Taken together, social media data could be useful for understanding the symptoms and epidemiology of novel diseases, such as COVID-19. Accumulated knowledge and techniques that utilize social media data would be rapidly applied for future pandemic preparedness and response in an effective manner.

Limitations and Future Work

The presence and rapid diffusion of misinformation in online communities is a growing concern for patients and health care providers [35]. We suspected that similar experiences and common interests among patients would make strong emotional connections in online disease communities, enabling us to extract relatively reliable information compared to nondisease communities. It has been shown that online disease communities are mainly composed of patients, family members of patients, and caregivers who are in similar situations and want to obtain accurate information by posting their real experience-based inquiries online [36]. In addition, patients are more likely to express emotions and share aspects of their life in an online community that they would not share face-to-face [37]. Such similar experiences, common interests, emotional connections, and empathy could develop and operate patient empowerment by creating and disseminating relevant information that would help them to understand their health conditions and to receive psychological support. Furthermore, disease communities develop practices that improve the quality of the information that peers exchange [38]. Since disease communities deal with life-related matters, peer interactions in online disease communities have found low levels of inaccuracy [39], and most false or misleading statements are rapidly corrected by participants [40]. In addition to patients' honest voluntary participation, we adopted an active learning method by collecting manually labeled Reddit posts and performed second-round manual inspections of labeled posts to improve the authenticity of the social media data. We believe these efforts have greatly improved the reliability and authenticity of our study.

It could be possible that all the symptoms extracted from Reddit would not cover the entire literature of COVID-19 symptoms or would not necessarily be COVID-19 symptoms. Moreover, there is an inherent uncertainty in social media analysis about the accuracy of temporal directions and the potential delays that might occur between the actual event and the time of posting, which is challenging to detect. A closely related limitation is our assumption of a duration of 20 days for the COVID-19 patient journey, which is derived through the available data (Reddit posts).

Natural language processing (NLP) could assist in identifying the temporal mentions and adjusting the outcomes. We might be able to improve our recovery time estimates using NLP, helping to improve temporal accuracy. As we work beyond social media surveillance and integrate with other data sets in the future, we are likely to find better alignment regarding the patient journey and related timelines.

We believe that the integration of various social media data sets and the accumulation of data or systematic surveys of COVID-19 survivors as well as customized machine learning algorithms to capture COVID-19 symptoms would help to decide which symptoms are manifestations of COVID-19 infection and provide new and untapped insights into understanding the longitudinal evolution of symptoms throughout the entire COVID-19 patient journey. This also could be a novel means of assessing the fraction of COVID-19 survivors with persistent symptoms.

Conclusion

In this observational study, we demonstrated the extensive variability of physiological and psychological impacts of COVID-19 infection and their variability during the acute infection and recovery phases of illness. We also demonstrated the usefulness of gathering social media data as an effective and alternative way to understanding the patient journey from diagnosis through recovery. Our findings show the practicality and feasibility of employing social media data for investigating disease states and understanding the evolution of physiological and psychological characteristics of disease over time. These practices could be incorporated into routine procedures for COVID-19 patient care, providing appropriate treatment and long-term care after recovery from COVID-19.

Acknowledgments

We would like to thank all members of the Applied Sciences team at Klick Inc for their helpful advice and assistance. JJ and GB designed the study. JJ, GB, and YF collected the data. JJ and SS analyzed and interpreted the data. JJ wrote the manuscript. Especially, we thank Jana Foley, Tae Wook Kang, Kinzie Earle, Nicholas Terpstra, Gai Paupaw, Paul Bourke, Aida Cagula, Dallas Gao, Sean Brice, Louisana Baldoz, Vlada Ross, Dave Coley, and Gaurav Baruah for assessing posts and assigning labels. All authors in this study revised the manuscript and contributed to the final review and editing and have approved the final manuscript. This research was internally funded and received no specific grant from any funding agency in the public, commercial, or not-for-profit sector.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary methods, figures, and tables.
[DOCX File , 4643 KB - jmir_v24i2e33959_app1.docx]

References

1. Guan W, Ni Z, Hu Y, Liang W, Ou C, He J, China Medical Treatment Expert Group for Covid-19. Clinical characteristics of coronavirus disease 2019 in China. N Engl J Med 2020 Apr 30;382(18):1708-1720 [FREE Full text] [doi: 10.1056/NEJMoa2002032] [Medline: 32109013]
2. Wu C, Chen X, Cai Y, Xia J, Zhou X, Xu S, et al. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. JAMA Intern Med 2020 Jul 01;180(7):934-943 [FREE Full text] [doi: 10.1001/jamainternmed.2020.0994] [Medline: 32167524]
3. Guo J, Radloff CL, Wawrzynski SE, Cloyes KG. Mining Twitter to explore the emergence of COVID-19 symptoms. Public Health Nurs 2020 Nov;37(6):934-940 [FREE Full text] [doi: 10.1111/phn.12809] [Medline: 32937679]
4. Sarker A, Lakamana S, Hogg-Bremer W, Xie A, Al-Garadi MA, Yang Y. Self-reported COVID-19 symptoms on Twitter: an analysis and a research resource. J Am Med Inform Assoc 2020 Aug 01;27(8):1310-1315 [FREE Full text] [doi: 10.1093/jamia/ocaa116] [Medline: 32620975]
5. Mizrahi B, Shilo S, Rossman H, Kalkstein N, Marcus K, Barer Y, et al. Longitudinal symptom dynamics of COVID-19 infection. Nat Commun 2020 Dec 04;11(1):6208 [FREE Full text] [doi: 10.1038/s41467-020-20053-y] [Medline: 33277494]
6. Jeon J, Baruah G, Sarabadani S, Palanica A. Identification of risk factors and symptoms of COVID-19: analysis of biomedical literature and social media data. J Med Internet Res 2020 Oct 02;22(10):e20509 [FREE Full text] [doi: 10.2196/20509] [Medline: 32936770]
7. Willi S, Lüthold R, Hunt A, Hänggi NV, Sejdiu D, Scaff C, et al. COVID-19 sequelae in adults aged less than 50 years: a systematic review. Travel Med Infect Dis 2021;40:101995 [FREE Full text] [doi: 10.1016/j.tmaid.2021.101995] [Medline: 33631340]
8. Nalbandian A, Sehgal K, Gupta A, Madhavan MV, McGroder C, Stevens JS, et al. Post-acute COVID-19 syndrome. Nat Med 2021 Apr;27(4):601-615. [doi: 10.1038/s41591-021-01283-z] [Medline: 33753937]
9. Nakamichi K, Shen JZ, Lee CS, Lee A, Roberts EA, Simonson PD, et al. Hospitalization and mortality associated with SARS-CoV-2 viral clades in COVID-19. Sci Rep 2021 Feb 26;11(1):4802 [FREE Full text] [doi: 10.1038/s41598-021-82850-9] [Medline: 33637820]
10. Hasan A, Levene M, Weston D. Learning structured medical information from social media. J Biomed Inform 2020 Oct;110:103568 [FREE Full text] [doi: 10.1016/j.jbi.2020.103568] [Medline: 32942027]
11. Kim J, Lee D, Park E. Machine learning for mental health in social media: bibliometric study. J Med Internet Res 2021 Mar 08;23(3):e24870 [FREE Full text] [doi: 10.2196/24870] [Medline: 33683209]
12. Baruah G, Zhang H, Guttikonda R, Lin J, Smucker M, Vechtomova V. Optimizing nugget annotations with active learning. 2016 Presented at: Proceedings of the 25th ACM International Conference On Information and Knowledge Management; 2016; Indianapolis, USA p. 2359-2364. [doi: 10.1145/2983323.2983694]
13. Rios P, Radhakrishnan A, Williams C, Ramkissoon N, Pham B, Cormack GV, et al. Preventing the transmission of COVID-19 and other coronaviruses in older adults aged 60 years and above living in long-term care: a rapid review. Syst Rev 2020 Sep 25;9(1):218 [FREE Full text] [doi: 10.1186/s13643-020-01486-4] [Medline: 32977848]
14. Wu X, Chen C, Zhong M, Wang J, Shi J. COVID-AL: the diagnosis of COVID-19 with deep active learning. Med Image Anal 2021 Feb;68:101913 [FREE Full text] [doi: 10.1016/j.media.2020.101913] [Medline: 33285482]
15. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Ann Intern Med 2020 May 05;172(9):577-582 [FREE Full text] [doi: 10.7326/M20-0504] [Medline: 32150748]

16. Bhapkar HR, Mahalle PN, Dey N, Santosh KC. Revisited COVID-19 mortality and recovery rates: are we missing recovery time period? J Med Syst 2020 Oct 25;44(12):202 [FREE Full text] [doi: 10.1007/s10916-020-01668-6] [Medline: 33099706]
17. Baud D, Qi X, Nielsen-Saines K, Musso D, Pomar L, Favre G. Real estimates of mortality following COVID-19 infection. Lancet Infect Dis 2020 Jul;20(7):773 [FREE Full text] [doi: 10.1016/S1473-3099(20)30195-X] [Medline: 32171390]
18. Singanayagam A, Patel M, Charlett A, Lopez Bernal J, Saliba V, Ellis J, et al. Duration of infectiousness and correlation with RT-PCR cycle threshold values in cases of COVID-19, England, January to May 2020. Euro Surveill 2020 Aug;25(32):2001483 [FREE Full text] [doi: 10.2807/1560-7917.ES.2020.25.32.2001483] [Medline: 32794447]
19. Voinsky I, Baristaite G, Gurwitz D. Effects of age and sex on recovery from COVID-19: analysis of 5769 Israeli patients. J Infect 2020 Aug;81(2):e102-e103 [FREE Full text] [doi: 10.1016/j.jinf.2020.05.026] [Medline: 32425274]
20. Mark N, Daniel K, Iz B, Waleed A. ScispaCy: fast and robust models for biomedical natural language processing. 2019 Presented at: Proceedings of the 18th BioNLP Workshop and Shared Task; 2019; Florence, Italy p. 319-327. [doi: 10.18653/v1/w19-5034]
21. Chengqi Z, Shichao Z. Association Rule Mining: Models and Algorithms. Berlin: Springer; 2002.
22. Parminder B, Busra C, Mohammed K, Selvan S. Comprehend Medical: a named entity recognition and relationship extraction web service. 2019 Presented at: 18th IEEE International Conference On Machine Learning And Applications (ICMLA); 2019; Boca Raton, FL p. a-51. [doi: 10.1109/icmla.2019.00297]
23. P'ng C, Green J, Chong LC, Waggott D, Prokopec SD, Shamsi M, et al. BPG: Seamless, automated and interactive visualization of scientific data. BMC Bioinform 2019 Jan 21;20(1):42 [FREE Full text] [doi: 10.1186/s12859-019-2610-2] [Medline: 30665349]
24. Tian S, Hu N, Lou J, Chen K, Kang X, Xiang Z, et al. Characteristics of COVID-19 infection in Beijing. J Infect 2020 Apr;80(4):401-406 [FREE Full text] [doi: 10.1016/j.jinf.2020.02.018] [Medline: 32112886]
25. Lin C, Zhang Y, Zhang K, Zheng Y, Lu L, Chang H, et al. Fever promotes T lymphocyte trafficking via a thermal sensory pathway involving heat shock protein 90 and α4 integrins. Immunity 2019 Jan 15;50(1):137-151.e6 [FREE Full text] [doi: 10.1016/j.immuni.2018.11.013] [Medline: 30650373]
26. Song W, Chang Y. Cough hypersensitivity as a neuro-immune interaction. Clin Transl Allergy 2015;5:24 [FREE Full text] [doi: 10.1186/s13601-015-0069-4] [Medline: 26180629]
27. Bartley JM, Pan SJ, Keilich SR, Hopkins JW, Al-Naggar IM, Kuchel GA, et al. Aging augments the impact of influenza respiratory tract infection on mobility impairments, muscle-localized inflammation, and muscle atrophy. Aging (Albany NY) 2016 Apr;8(4):620-635 [FREE Full text] [doi: 10.18632/aging.100882] [Medline: 26856410]
28. Rathor R, Suryakumar G, Singh SN. Diet and redox state in maintaining skeletal muscle health and performance at high altitude. Free Radic Biol Med 2021 Oct;174:305-320. [doi: 10.1016/j.freeradbiomed.2021.07.024] [Medline: 34352371]
29. Chan AKM, Nickson CP, Rudolph JW, Lee A, Joynt GM. Social media for rapid knowledge dissemination: early experience from the COVID-19 pandemic. Anaesthesia 2020 Dec 30;75(12):1579-1582 [FREE Full text] [doi: 10.1111/anae.15057] [Medline: 32227594]
30. Picone M, Inoue S, DeFelice C, Naujokas MF, Sinrod J, Cruz VA, et al. Social listening as a rapid approach to collecting and analyzing COVID-19 symptoms and disease natural histories reported by large numbers of individuals. Popul Health Manag 2020 Oct 01;23(5):350-360. [doi: 10.1089/pop.2020.0189] [Medline: 32897820]
31. Symptoms of COVID-19. National Center for Immunization and Respiratory Diseases (NCIRD), Division of Viral Diseases. URL: https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html [accessed 2022-02-07]
32. Yousefinaghani S, Dara R, Mubareka S, Sharif S. Prediction of COVID-19 waves using social media and google search: a case study of the US and Canada. Front Public Health 2021;9:656635 [FREE Full text] [doi: 10.3389/fpubh.2021.656635] [Medline: 33937179]
33. Shen C, Chen A, Luo C, Zhang J, Feng B, Liao W. Using reports of symptoms and diagnoses on social media to predict COVID-19 case counts in Mainland China: observational infoveillance study. J Med Internet Res 2020 May 28;22(5):e19421 [FREE Full text] [doi: 10.2196/19421] [Medline: 32452804]
34. Park Y. Developing a COVID-19 crisis management strategy using news media and social media in big data analytics. Soc Sci Comput Rev 2021 Apr 21:089443932110073. [doi: 10.1177/08944393211007314]
35. Domínguez M, Sapiña L. Pediatric cancer and the internet: exploring the gap in doctor-parents communication. J Cancer Educ 2015 Mar;30(1):145-151. [doi: 10.1007/s13187-014-0700-4] [Medline: 24986308]
36. Oh J, Kim JA. Information-seeking behavior and information needs in patients with amyotrophic lateral sclerosis: analyzing an online patient community. Comput Inform Nurs 2017 Jul;35(7):345-351. [doi: 10.1097/CIN.0000000000000333] [Medline: 28225381]
37. Hargreaves S, Bath PA, Duffin S, Ellis J. Sharing and empathy in digital spaces: qualitative study of online health forums for breast cancer and motor neuron disease (amyotrophic lateral sclerosis). J Med Internet Res 2018 Jun 14;20(6):e222 [FREE Full text] [doi: 10.2196/jmir.9709] [Medline: 29903695]
38. Hartzler A, Huh J. Level 3: patient power on the web: the multifaceted role of personal health wisdom. In: Wetter T, editor. Consumer Health Informatics: New Services, Roles, and Responsibilities. Cham: Springer; 2016:134-146.

39. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. BMJ 2004 May 15;328(7449):1166 [FREE Full text] [doi: 10.1136/bmj.328.7449.1166] [Medline: 15142921]
40. Esquivel A, Meric-Bernstam F, Bernstam EV. Accuracy and self correction of information received from an internet breast cancer list: content analysis. BMJ 2006 Apr 22;332(7547):939-942 [FREE Full text] [doi: 10.1136/bmj.38753.524201.7C] [Medline: 16513686]

Abbreviations

API: Application Programming Interface
NLP: natural language processing

Edited by A Mavragani; submitted 30.09.21; peer-reviewed by A Hasan, N Hu; comments to author 22.10.21; revised version received 11.11.21; accepted 22.01.22; published 16.02.22.

Please cite as:
Sarabadani S, Baruah G, Fossat Y, Jeon J
Longitudinal Changes of COVID-19 Symptoms in Social Media: Observational Study
J Med Internet Res 2022;24(2):e33959
URL: https://www.jmir.org/2022/2/e33959
doi:10.2196/33959
PMID:35076400

Original Paper

Concerns Around Opposition to the Green Pass in Italy: Social Listening Analysis by Using a Mixed Methods Approach

Giovanni Spitale[1], MA; Nikola Biller-Andorno[1], MD, PhD; Federico Germani[1], PhD

Institute of Biomedical Ethics and History of Medicine, University of Zurich, Zurich, Switzerland

Corresponding Author:
Federico Germani, PhD
Institute of Biomedical Ethics and History of Medicine
University of Zurich
Winterthurerstrasse 30
Zurich, 8006
Switzerland
Phone: 1 44 634 40 14
Email: federico.germani@ibme.uzh.ch

Abstract

Background: The recent introduction of COVID-19 certificates in several countries, including the introduction of the European green pass, has been met with protests and concerns by a fraction of the population. In Italy, the green pass has been used as a nudging measure to incentivize vaccinations because a valid green pass is needed to enter restaurants, bars, museums, or stadiums. As of December 2021, a valid green pass can be obtained by being fully vaccinated with an approved vaccine, recovered from COVID-19, or tested. However, a green pass obtained with a test has a short validity (48 hours for the rapid test, 72 hours for the polymerase chain reaction test) and does not allow access to several indoor public places.

Objective: This study aims to understand and describe the concerns of individuals opposed to the green pass in Italy, the main arguments of their discussions, and their characterization.

Methods: We collected data from Telegram chats and analyzed the arguments and concerns that were raised by the users by using a mixed methods approach.

Results: Most individuals opposing the green pass share antivaccine views, but doubts and concerns about vaccines are generally not among the arguments raised to oppose the green pass. Instead, the discussion revolves around the legal aspects and the definition of personal freedom. We explain the differences and similarities between antivaccine and anti–green pass discourses, and we discuss the ethical ramifications of our research, focusing on the use of Telegram chats as a social listening tool for public health.

Conclusions: A large portion of individuals opposed to the green pass share antivaccine views. We suggest public health and political institutions to provide a legal explanation and a context for the use of the green pass, as well as to continue focusing on vaccine communication to inform vaccine-hesitant individuals. Further work is needed to define a consensual ethical framework for social listening for public health.

(J Med Internet Res 2022;24(2):e34385) doi:10.2196/34385

KEYWORDS

green pass; COVID-19; COVID-19 pandemic; vaccines; vaccination hesitancy; freedom; social listening; social media; infodemic; bioethics; telegram

Introduction

Background

Since the beginning of large-scale vaccination campaigns against COVID-19, many countries have had to deal with the issue of vaccine hesitancy [1]. Already defined by the World Health Organization in 2019 as one of the major threats to global health [2], vaccine hesitancy has become even more relevant in the context of the COVID-19 pandemic [3]. In Israel, the first country that was able to ensure sufficient supplies of the Pfizer-BioNTech vaccine, the Ministry of Health swiftly started a vaccination campaign in late 2020. However, after covering health care staff, older adults, and vulnerable patients, the campaign reached a stagnation phase, as a relevant percentage of individuals were not willing to get vaccinated. After

considering other forms of incentives [4], the Israeli Ministry of Health developed a new ad hoc strategy to increase the vaccination rate. According to this plan, vaccinated people would receive a special document that would allow them access to social and cultural events, national and international mobility, and exemption from quarantine. The declared aim of this document or the "green passport" was to encourage citizens to receive COVID-19 vaccinations while allowing some reopening of the economy [5]. The proposal for the Israeli green passport was passed on December 14, 2020 [4]; on January 27, 2021, the eHealth network of the European Commission started to develop a set of guidelines in order to implement a "EU Digital COVID certificate" system in Europe. On June 1, 2021, the European Union Gateway, that is, the backbone interconnecting national green pass systems in the European Union, went live [6]. On July 23, 2021 the Italian government passed Decree-Law 105, regulating the use of the green pass [7], already recognized and defined on April 22, 2021 by Decree-Law 52, article 9 [8].

Compared to other nudging strategies to tackle vaccine hesitancy, the green pass looks like a promising concept, as it incentivizes people to get vaccinated without imposing a decision; however, already in its first implementation in Israel, it generated some debate as it can be considered as a tool for discrimination based on someone's vaccination status. Another argument often used by green pass critics regards privacy; when showing their green pass, people are de facto obliged to disclose health information—thus, their sensitive information—to third parties [5]. The adoption of the green pass strategy in Europe caused the very same debate and the very same arguments already seen in Israel. In many European countries, opposition movements started to form and grow, discussing the use of the green pass and organizing protests, sit-ins, and rallies [9]. In a time of physical distancing due to containment measures, many of these discussions were taking place online on social media and communication platforms. As popular social media platforms increasingly corrected their policies to decrease the flow of misinformation [10,11], people and organizations holding critical views about the green pass started to deplatform toward alternative social media channels, a phenomenon already seen and studied, mostly regarding the far-right and conservative world [12]. Notably, one of the most prominent destinations for deplatformed individuals and organizations has been Telegram.

Over the last few years, Telegram has become one of the most prominent instant messaging services. This success is due to a combination of 2 factors: on the one hand, end-to-end encryption [13] and an infrastructure distributed over several jurisdictions [14] makes it rather difficult to extract data from the system [15]. As stated on the official Telegram's frequently asked questions, to this day, Telegram has "disclosed 0 bytes of user data to third parties, including governments" [14]. On the other hand, Telegram's services go way beyond conventional instant messaging services: Telegram groups allow a maximum of 200,000 members and include advanced features such as unified history, instant search, replies, permissions, and moderation tools, making them outstanding tools for many-to-many discussions. In parallel, Telegram broadcast channels allow an unlimited number of followers, making them an appealing alternative to Twitter for one-to-many communication [14,16]. This combination of publicity, mobilization capabilities, and privacy provides a solution to the so-called "terrorists' dilemma," that is, the balancing security and outreach in choosing a web-based communication platform [17]. The use of Telegram among "no-green-pass groups" in Italy started to grow rapidly already in July 2021; as soon as the green pass was introduced, groups and individuals offering forged green passes for purchase started to exist [18] as well as groups organizing protests and rallies against the green pass [19].

Aims

This study has 2 aims: (1) to study the discourse revolving around the opposition to the green pass and its use in Telegram chats by no-green-pass groups in Italy, with a focus on groups used by university students; and (2) to detail a novel approach to online social listening by using a combination of quantitative and qualitative approaches and to question its ethical aspects.

Methods

Ethical and Legal Considerations

As this study did not fall under the scope of the Swiss Human Research Act [20], authorization from the Cantonal Ethics Committee was not required. The messages analyzed in this study were retrieved from public chats using the "download history" function of Telegram Desktop. This qualifies the data as publicly available. According to the General Data Protection Regulation (GDPR) [21] article 6.1, data processing without explicit consent of data subjects is possible when protecting the interest of the data subject and when "necessary for the performance of a task carried out in the public interest." Research falls in the category of public interest, but this criterion being very broad, it is important to weigh the public interest and benefits to the risk for the individuals, especially because the data set might contain special categories of personal data (ie, health, politics, or worldview-related data). Generally, the information detailed in article 14 of the GDPR should be provided to the data subjects individually, although this could be considered as a disproportionate effort, given the number of users involved in this study. However, one could argue that the necessary information could be provided in a general way through posting into those chats. Since either way this transparency might result in both a higher risk of reidentification and a serious impairment to the pursue of research, it could be argued that is against the public interest and should therefore be omitted. Articles 14.5 and 89 of the GDPR exempt from the provision of information to study participants where and insofar it would involve a disproportionate effort or render impossible or seriously impair the achievement of the objectives (ie, the research goals in the public interest). As specified in article 14.5.b, we took appropriate measures to protect the privacy of data subjects whose messages are included in our study: the JSON files retrieved from Telegram have been completely anonymized (removal of personal names and toponyms from the message text) and pseudonymized (replacement of the user ID with a pseudonym), the original data set has been destroyed, the analysis has been conducted on the anonymized version, the anonymized data set will be available upon request, and as

the search of segments of text in the original chat would allow reidentification, the links to the chats will not be disclosed.

Data Collection

Data were collected from 2 groups of chats. The first comprised no-green-pass groups of Italian universities (one in the north, one in the center, and one in the south) and generic no-green-pass groups. The second, our negative control, comprised groups discussing topics unrelated to COVID-19 or to the green pass, such as video games, parrot breeding, and other general topics. The selection of these control chats followed 3 criteria: the chats were active (at least 10 messages sent in the last week), they counted at least 200 users, and the main language used was Italian. We identified relevant chats and downloaded the message history as JSON files. We downloaded the JSON files containing the entire history of the said groups on September 9, 2021. The data collection is described in Multimedia Appendix 1.

Data were downloaded directly using the function "export chat history" of Telegram's official desktop client. For this study, we downloaded only textual data. We parsed the JSON files into Pandas data frames. To protect the privacy of the users while still maintaining the possibility to track conversations in qualitative analysis, we combined anonymization and pseudonymization. Anonymization was performed by removing metadata from messages and by removing personal names and replacing them with [name] [22]. Similarly, every toponym was replaced with [place] [23]. Direct mentions of users in the text (eg, @thisuser) were searched and replaced with [username]. Surnames were not removed from messages. Since the chats were in an informal context, people did not refer to other members of the chat or to themselves using surnames. However, surnames are often used to refer to public figures or sources of information and thus represent a valuable component for the analysis.

Analysis

For this project, first, we used a mixed methods approach, which involves the use of qualitative and quantitative data. Second, for the quantitative analysis, with a top-down approach, we defined a series of dictionaries relevant for the purpose of this study—each one containing regular expressions that belong to the same concept. Regex (Regular Expressions) allows the definition of fairly complex rules, able to reduce ambiguity, and capture precise concepts. As an example, the rule (tesser.\sverd.?|pass\sverd.?|certifica\w*\sverd.?) will fire on "tessera verde" (green pass) or "tessere verdi" (green passes) or "pass verde" (green pass) or "certificato verde" (green certificate) but not on "casa verde" (green house) or "verderame" (verdigris) or "tessera del cinema" (cinema card). The autocoding has a weight system: if only one rule from the dictionary fires, the autocode is assigned a weight of 1, if 2 rules fire, the weight will be 2 and so on. Autocodes can then be used to measure the prevalence of topics through the corpus to segment the quantitative analyses or as a starting point for the qualitative work. Third, we extracted the lemmas used in the corpus by using the Python package "spaCy" and its pretrained model for Italian [24]. This was performed on a large bag of words including every message in the corpus and by dividing messages by code. In the final step of the quantitative analysis, we performed a sentiment analysis [25], both on the entire corpus and on messages divided by code. The sentiment analysis was performed using the Python package "feel-it" [26], through which we calculated the probability of positive or negative sentiment for each message. We developed the analysis pipeline in Python; the code is structured in a JupyterLab notebook, available through Zenodo [27].

For the qualitative analysis, we generated a structured text file, annotated with pseudonymized speakers and codes resulting from the autocoding system. The file was then imported into MAXQDA for thematic analysis. The development of the regular expressions used for autocoding has been an iterative process. We ran the code several times, exploring the results, noting the gaps, and fine-tuning the regular expressions. The thematic analysis has been conducted by native Italian speakers on messages written in Italian; the text has been translated by the authors to be comprehensible to a wider audience but still as close as possible to the original. The original quotes in Italian and the categories/topics included in the analysis are provided as supplementary material (Multimedia Appendix 2 and Multimedia Appendix 3).

Results

Quantitative Results

Lemmas, Terms, and Rules: the No-Green-Pass Discourse Encompasses Legal Aspects, Actions, and Vaccine Skepticism

To understand the interests of individuals critical of the green pass, their arguments, and the opinions that shape their position in the debate, we quantified and analyzed the most frequently used lemmas in control chats (Multimedia Appendix 4)—discussing issues not related to green pass, vaccines, or COVID-19, and in chats focused on green pass opposition (Multimedia Appendix 5). As a positive control, we checked whether the lemmas "green" and "pass" were found to be among the most frequently used in green pass opposition chats when compared with control chats. As expected, "green" was the second most frequently used lemma in green pass opposition chats, and "pass" was the fourth most frequently used lemma (frequencies 9.2% and 7.5%, respectively). However, these lemmas were barely used in control chats (frequency of 0.02% for both lemmas). As expected, the average frequency of the 2 lemmas combined ("green" + "pass") in green pass opposition chats was significantly higher than that in control chats (Figure 1A). Among the 20 most used lemmas in either control or green pass opposition chats, we identified 2 relevant categories of terms: legal terms and action terms. Legal terms included law (*legge*) and article (*articolo*). These terms were highly overrepresented in green pass opposition chats when compared with those in control chats (Figure 1B). Action lemmas included can (*potere*), must (*dovere*), want (*volere*), know (*sapere*), ask (*chiedere*), do (*fare*), say (*dire*), speak (*parlare*), take (*prendere*), put (*mettere*), use (*utilizzare*), come and go (*andare* and *venire*), and write (*scrivere*). Among these, we identified 3 lemmas to be relevant and underrepresented in green pass opposition chats: take (*prendere*), put (*mettere*), and use (*utilizzare*).

Overrepresented lemmas were can (*potere*), ask (*chiedere*), and speak (*parlare*) (Figure 1C). Besides these lemmas, we were interested in understanding which rules were the most relevant in control chats, and, in addition to the rule "green pass," we focused on the rules "COVID-19" and "vaccine" and "freedom." With this analysis, we wanted to understand whether the legal lemmas that were overrepresented in green pass opposition chats were linked to a pronounced discussion about personal freedom, in connection with the discussion about the green pass and COVID-19 vaccines. All the abovementioned rules were represented with a higher frequency in green pass opposition chats when compared with those in control chats, and these differences were statistically significant (Figure 2). As expected, the rule for "green pass" fired very frequently in green pass opposition chats and more frequently than the rules "COVID-19" and "freedom." Surprisingly, however, the rule "vaccine" was the most frequently used in green pass opposition chats, more so than the rule "green pass," indicating that among green pass critics, even when the discussion revolves around legal aspects connected to personal freedom, skepticism toward vaccines likely remains as the predominant reason to oppose the green pass.

Figure 1. (A) Average lemma frequency (in percentage) in control versus green pass opposition chats. Average lemma frequency (in percentage) for green pass in control (grey bar) versus green pass opposition chats (black bar). (B) Average lemma frequency (in percentage) for legal terms in control chats (grey bars) when compared with green pass opposition chats (black bars), extracted from the 20 most frequently used words in the green pass opposition chats. (C) Average lemma frequency (in percentage) for action terms in control (grey bars) versus green pass opposition chats (black bars), extracted from the 20 most frequently used words in both control and green pass opposition chats. The green background highlights the most relevant action terms that are overrepresented in the green pass opposition chats, whereas the grey background highlights the most relevant action terms that are overrepresented in control chats. **P*<.05, *t* test. Error bars represent standard error of the mean.

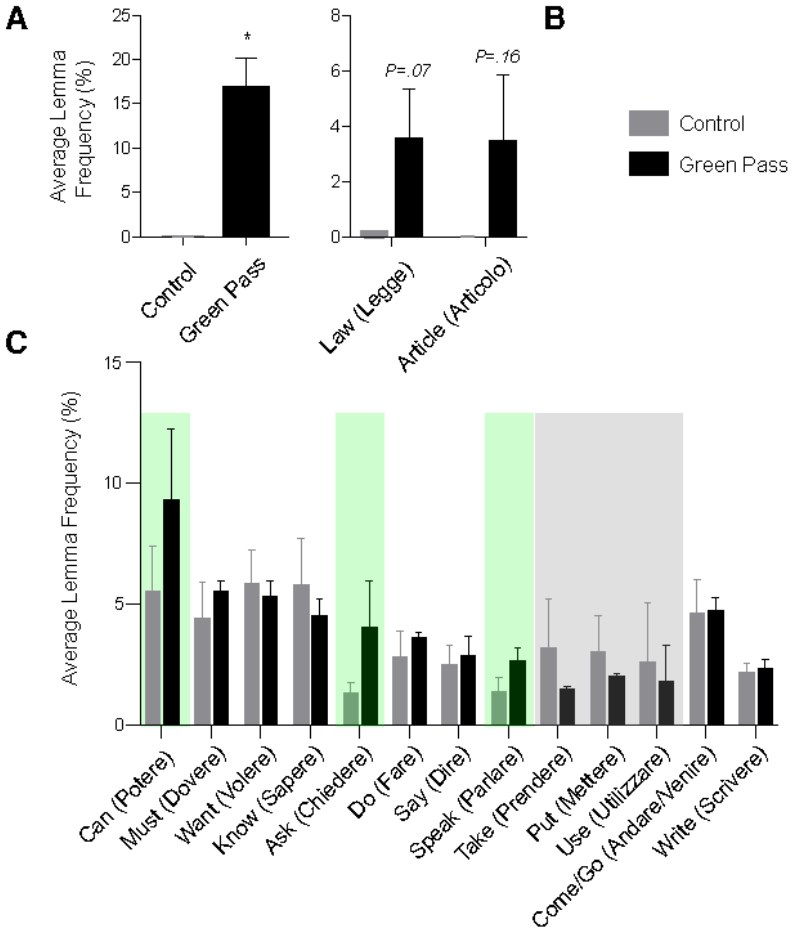
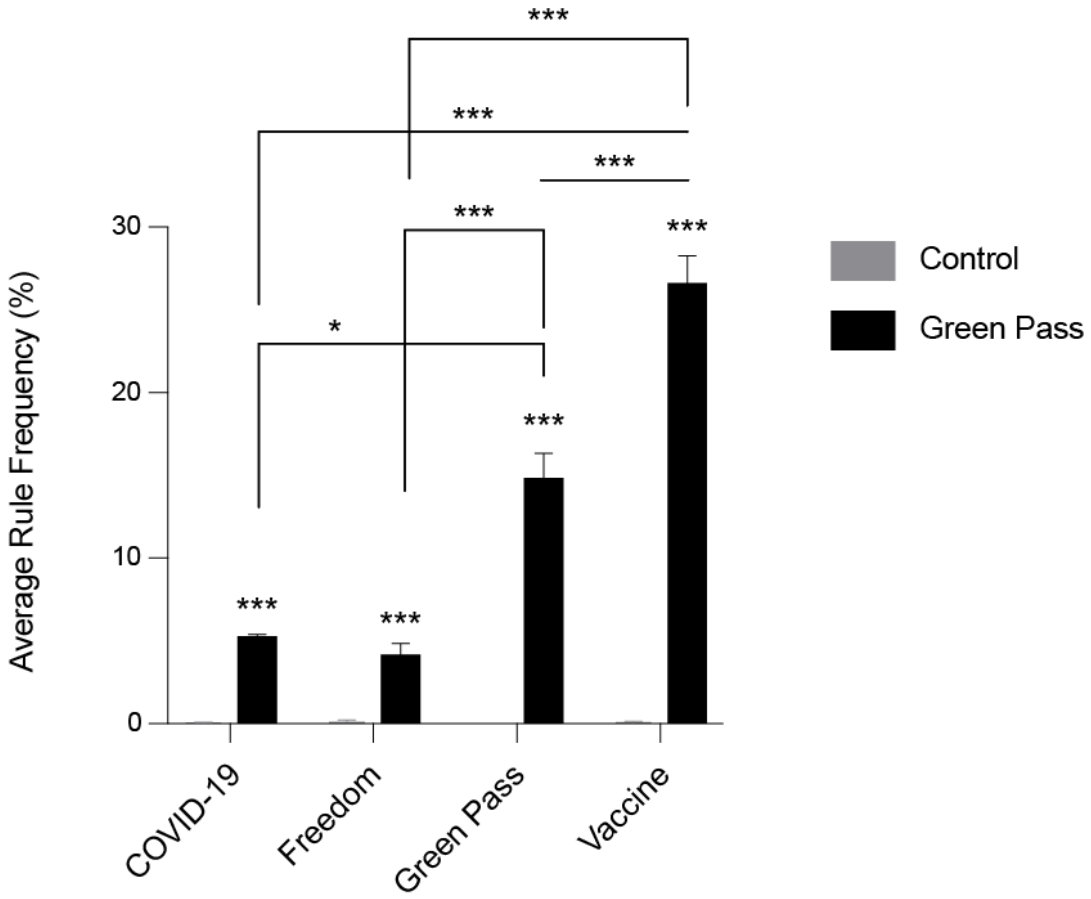


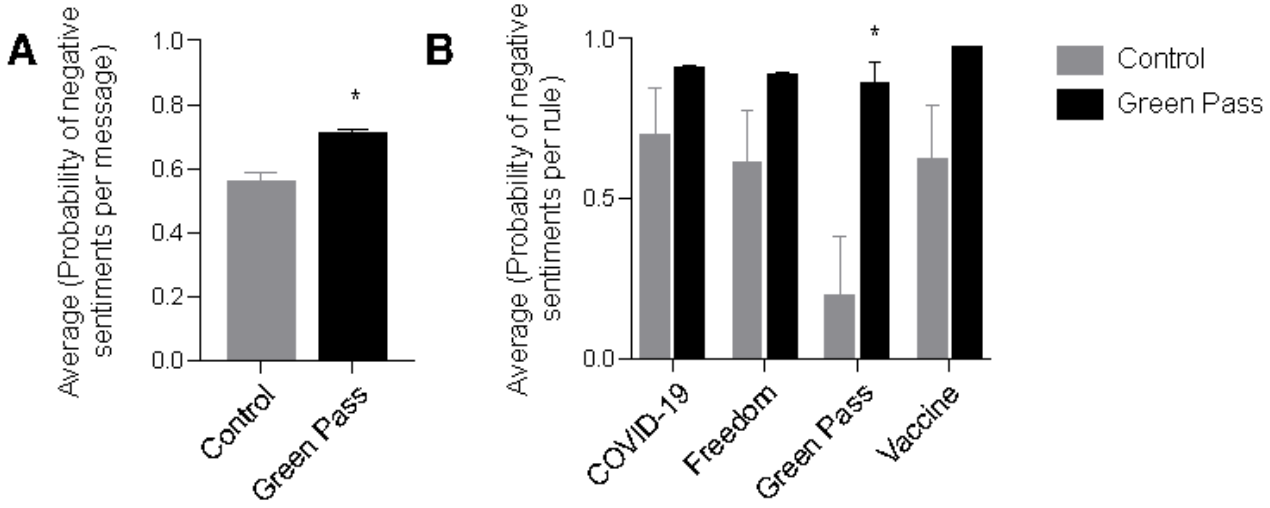
Figure 2. Average rule frequency (in percentage) in control versus green pass opposition chats. Average rule frequency for terms grouped under the rules “COVID-19,” “freedom,” “green pass,” and “vaccine” in control (grey bars) versus green pass opposition chats (black bars). *P<.05; ***P<.001, t test. Error bars represent standard error of the mean.



“No-Green-Pass” Individuals Have a Negative Sentiment Toward Green Pass and Vaccines

After having identified the predominant themes associated with anti–green pass discourse, we analyzed whether such a discourse is associated with a higher probability of negative sentiments. By defining the likelihood of negative sentiment for each message, we averaged the sentiment for each chat and finally across chats within the same category. As expected, the average likelihood of negative sentiment was significantly higher in green pass opposition chats when compared to that in control chats, with a probability of 0.70 and 0.55, respectively (Figure 3A). In addition, we calculated the average probability of negative sentiment associated with the rules “COVID-19,” “freedom,” “green pass,” “vaccine,” and determined that for all these rules, messages depicting negativity were overrepresented in green pass opposition chats when compared with those in control chats (Figure 3B). This effect was significant for the rule “green pass,” which can serve as a positive control, indicating that green pass critics are, in fact, assessing the issue with negative sentiment, when compared with people that do not necessarily oppose its introduction and use. Of particular interest, messages related to the rule "vaccine" had a 96.26% probability to depict negative sentiment, a particularly high probability also when compared with negativity for "COVID-19", "freedom", and "green pass" in green pass opposition chats (90%, 88%, and 85%, respectively), thus providing strength to the hypothesis that vaccine skepticism is the primary reason to oppose the green pass.

Figure 3. Sentiment analysis in control versus green pass opposition chats. Average probability of negative sentiment in messages published in control (grey bar) versus green pass opposition (black bar) chats. (A) Average probability of negative sentiment per rule in control (grey bars) versus green pass opposition chats (black bars). The following rules are taken into consideration: "COVID-19," "Freedom," "Green Pass," and "Vaccine" (B) 0 indicates the maximum likelihood for an average message to display positive sentiment, whereas 1 indicates the maximum likelihood for an average message to display negative sentiment. *P<.05, t test. Error bars represent standard error of the mean.



Rules and Lemma Frequency: Interplay Between Vaccines and Green Pass

To further understand the relationship among the topics "green pass," "vaccine," "freedom," and "COVID-19," we analyzed the most frequently used lemmas when the discussion was on one of such topics (as determined using the associated rules). For the rule "COVID-19," the lemmas "green," pass," and "vaccine" were among the most used (Figure S1A in Multimedia Appendix 6). For the rule "freedom," as expected, lemmas associated with legal terms were overrepresented as well as "green" and "pass" (Figure S1B in Multimedia Appendix 6). For the term "green pass," we could not identify "vaccine" among the most relevant and used lemmas, although we identified lemmas associated with legal terms, including "freedom," "law," "article," "can," and "must" (Figure S1C in Multimedia Appendix 6). Instead, for the rule "vaccine," we could identify "green" and "pass" among the most relevant and significant lemmas (Figure S1D in Multimedia Appendix 6). As our previous results indicate, although our analysis was focused on green pass opposition chats, vaccines constituted a widely discussed topic, which even dominated the discussion about the green pass. In line with our previous observations, here we show that green pass discussion takes place when vaccines are being discussed but not vice versa. This might suggest that green pass critics tend to share antivaccine views but do not wish their argumentations against the green pass to revolve around their antivaccine views. Rather, they prefer to support their position by discussing limitations to personal freedom and advancing legal considerations.

Qualitative Results

Green Pass and Vaccines

The qualitative analysis supports the findings described in our quantitative analysis. Although our analysis was focused on chats discussing the green pass, users often started debating about related topics, including the risk-benefit profile of COVID-19 vaccines, their efficacy, and their use. Of note, moderators often asked participants to stay on topic and avoid discussing these parallel issues. There were 2 main reasons: one was to avoid conflict, as a (small) fraction of individuals who positioned themselves as opposed to the green pass were for vaccinations; the other was to avoid floods of misinformation, which could discredit what the moderators perceived as a much needed debate. The quotes below are shown by the subcorpus (university group or generic group, and the position refers to the line in the subcorpus).

> *…On the other hand, it is a big mistake to take a stance on vaccines. Those who want to do so should do so. The point is only to be against this limitation of freedom and many vaccinated people are against the green pass. Do not introduce divisive or extremist elements that vote the initiative down.* [University, south, position 742]

Users themselves are very aware of how hard it is to discuss the green pass without discussing the reasons for which it is needed.

> *…how can one ignore the vaccine issue if it is literally the main option for getting a pass?* [University, north, position 6693]

As stated above and as noted in the quantitative analysis, these no-green-pass chats have de facto been a proxy to discuss vaccines. Users know the green pass was introduced as a nudging measure to avoid mandatory vaccination. Nevertheless, they do not perceive this strategy in a positive way. Even if there are other ways to receive a valid green pass (ie, recovery from COVID-19 or testing), vaccination is the most obvious and less burdensome one. Users perceive this as a cunning imposition, which possibly makes them even angrier than mandated vaccines.

> *…I am against the green pass because I see it as a coercive and hypocritical tool put in place by the government because if they saw the vaccine as a safe*

> *way, they should have the consistency to make it compulsory and instead they don't bother to do so.* [University, south, position 1807]

> *…the green pass is a way of circumventing compulsory vaccination. The green pass is an "incentive," said to be very soft, but in fact it is a compulsory requirement.* [University, center, position 14716-14718]

Even though, as discussed above, moderators would prefer to disentangle the discussion about the green pass from the topic of vaccinations, pointing out that even someone who is vaccinated could hold no-green-pass positions, most of the users shared common critical beliefs about vaccines.

> *…It is becoming increasingly clear from the scientific literature that (1) There are very effective treatments for COVID that indicate that vaccines are not at all essential. (2) Vaccines often have serious short-, medium- and long-term side effects, there is a well-founded fear that they could induce serious pathologies (tumors, autoimmune and degenerative diseases, sterility...) and they are still at the experimental stage. (3) Vaccines facilitate the development of variants, many of which are particularly virulent, and should not be carried out during epidemics, let alone pandemics. (4) Vaccines do not absolutely protect against COVID as they are said to do, ie, those vaccinated may become infected and may in turn infect others... so they should not have a Green pass unless they too are swabbed.* [University, center, position 3572-3579]

Users are especially afraid of the possible side effects. This narrative proposes that the vaccine is worse than the disease it is meant to prevent.

> *…who can guarantee that I will not have serious effects as a result of the vaccine, which could harm my future? Who will compensate me for any damage?* [University, north, position 25293-25294]

> *…Statistics show that the number of deaths due to COVID is the same as the number of deaths due to the vaccine, only that the number of deaths due to COVID is much overestimated (the number also includes deaths due to other causes but catalogued as COVID because they are positive to the test), while deaths due to the vaccine (not to mention cases of serious adverse effects) are much underestimated because only passive surveillance is done, and poorly.* [University, center, position 15682-15688]

Moreover, according to several users, there is no evidence that vaccines work. They do not prevent the spread of the disease and are less effective and more burdensome than alternative therapies to reduce the symptoms (the most quoted are hydroxychloroquine, cortisone, heparin, ivermectin, nonsteroidal anti-inflammatory drug, and hyperimmune plasma transfusions).

> *…We must rebel, this vaccine is a gene therapy with no guarantee that it will work. Vaccinated people are just as infectious as unvaccinated people, it is clear that this vaccine does not protect against COVID.* [University, north, position 2612]

> *…It is written in all official documents of the pharmaceutical companies and the WHO that there is no evidence that vaccination will stop the spread of the virus.* [University, north, position 3385]

Finally, some users suggested that vaccines could be part of a bigger scheme, again orchestrated by governments and covert powers, possibly aiming to reduce the world's population.

> *…Overpopulation, they have been saying this for years, and the Vax in my opinion serves to solve that problem, not COVID.* [University, south, position 2343]

> *…Their aim is to manipulate human beings by injecting them with a serum containing graphene, which can react with certain frequencies and modify the behavior of cells. By changing the behavior of cells, you can change the behavior of human beings.* [Generic, position 72471]

Beyond Vaccines: Green Pass, Legal Aspects, and Personal Freedom

Despite vaccines being the predominant topic in these chats, the majority of individuals did not make use of arguments related to vaccines, including conspiracy theories about vaccines, to justify their opposition to the green pass. Rather, they claimed the green pass was an illegal measure and it is discriminatory.

> *…it is clear that the green pass is an instrument of political discrimination that has no relation to the actual health status.* [University, center, position 3572-3579]

> *…The green pass is clearly unconstitutional and discriminatory in nature and is a purely political instrument as it has no scientific basis; the report linked before is very clear about it, then they do not make it mandatory by law otherwise they would be obliged to compensate those who died of the vaccine.* [University, center, position 7520-7522]

In some circumstances, users alluded to conspiracy theories according to which the green pass is an element of a bigger plan put in place either by governments or by covert powers to achieve other ends, usually the institution of a totalitarian regime.

> *…Do you still have to realize that even if the Regime decides to withdraw the COVID PASS, to let you go back to work, you have already become citizens of a totalitarian Regime? Citizens of a lousy Regime based on lies, on the progressive elimination of freedoms, on the violent suppression of dissent?* [Generic, position 2127]

Many users believed the green pass was a serious limitation of personal freedom. This argument was developed following 3 main threads in order of importance: jurisprudential, consequentialist, and deontological. On the jurisprudential side, users appeal mostly to the Italian constitution (articles 13 and

120), to law 196/2003 (personal data protection code), and to the Oviedo convention.

> *…the "green pass" cannot be checked because it is discriminatory, prejudicial to privacy and violates the following articles of law: - Art. 187 of the TULPS Regulation: a commercial operator is obliged to welcome in his business any person, without discrimination, under penalty of a fine up to €3000,00.- Privacy Law: no one can force us to provide information about our health conditions.- Art. 120 Italian Constitution: no one can limit the freedom of movement of the individual in the territory of the Italian republic. - Art. 13 Italian Constitution: no one may restrict personal freedom without a provision of the Judicial Authority on facts concerning the individual.* [Generic, position 3448]

> *…Add that we will respect all the anti-COVID security measures (social distancing, hygiene, mask). With regard to the reference to laws and treaties, don't we want to mention the convention on human rights, the Oviedo treaty and the Supreme Court ruling stating that the health of the individual cannot be sacrificed for the sake of collective health? don't we want to mention the principle of self-determination?* [University, center, position 395-397]

On the consequentialist side, users tend to fall into a slippery slope fallacy. In their view, the green pass system will necessarily lead to a system of capillary social control, repression of dissent, and loss of critical thinking capabilities—a system that is clearly undesirable and immoral; therefore, the green pass is undesirable and immoral as well.

> *…Look at the Chinese social score system to understand the crazy direction of these actions, typical of dictatorial systems and not of advanced democracies.* [University, south, position 3755]

> *…By now I think these people are lobotomized and probably don't even know the word FREEDOM.* [University, south, position 1255]

A minority of users tried to build a deontological argument, balancing values such as freedom and life. Their conclusion was that life and freedom have equal importance; hence, it is unjust to protect life by limiting freedom.

> *…If the answer to the question is that life is more important than liberty, then all the liberticidal laws made so far are justifiable and I would say almost fair, I can also understand why the green pass, a blatantly discriminatory law, is considered fair by many.*

> *…If the answer to the question is freedom, it is clear that everything that has been done so far is considered a mistake regardless of whether a particular law was made to save lives.*

> *…We come to the last answer, the most balanced one for me at least, that life and freedom are of equal importance. In view of this answer, it is clear that taking precautions to limit contagion and death is right and proper, so limitations will be inevitable (such as social distancing, masks indoors, limiting seating etc.), but at the same time it is important to preserve the freedoms of all citizens.* [University, center, position 14996-15012]

The 3 arguments converge on a single conclusion: the green pass and the system of control it creates are either seen as tools in the hands of dictators or as preparatory tools to gather power.

> *…What kind of disgusting nightmare do we want to bequeath to our children? A Health Regime? A Regime that brutalizes the minds and bodies of its citizens on a daily basis? Enough! Rebel!* [Generic, position 2127]

> *…We are living in a health dictatorship and political authoritarianism that must be opposed. I wonder if a general acting as a commissioner who comes out with absurd words about wanting to flush out the 'unvaccinated' house by house? These people must leave the government. We must demand to go to the polls again.* [University, center, position 5904-5906]

Action Plan

Leveraging on this understanding of freedom, users perceive a clear duty to react. The first and foremost action is understanding who the real enemy is, that is, not the virus, nor the people who get vaccinated or obey the regime. The real enemies are the political system and the political representatives who allowed this to happen.

> *…It is a political issue everywhere. If we understand this we know who we have to fight, and for sure it's not a virus.* [University, north, position 20112]

In the university groups, users discussed a lot about communication strategies that would allow them to be credible, also because they are aware that their groups might be studied. The most important points were regarding avoiding defusing topics (ie, conspiracy theories) and focusing on self-determination. Again, coherently with our quantitative findings, the main issue appeared to be the vaccine, for which the green pass was just a proxy.

> *…we have been able to ascertain the intense doxing activity also of Telegram groups. In short, now that membership is growing, we need a minimum of 'art of war' (or rather strategy, just to avoid accusations of terrorism).* [University, north, position 20233]

> *…no disquisitions that go beyond the topic to be defended, such as the existence or nonexistence of the virus, the no-pro vax diatribe, the Davis forum, depopulation, mass experimentation, variants, damage, etc. These are all topics on which one has burnt the candle at the stake. These are all topics on which the authoritativeness of many prominent figures has been burned, since they easily fall under the so to speak 'defusing' labels (conspiracist, degree obtained on Google, no Mask, no vax, no test, denialist).* [University, north, position 3607]

Lastly, many users considered protests as valid strategic options to make their voices heard. The options they considered ranged from flash mobs to general strikes, to occupations of the parliament.

> *…Shall we make a flash mob where all the unvaccinated all go in at the same time where they can't? Maybe running so that we are sweaty (so they are afraid to touch us) maybe with a hat that says "the Jew rebels"* [Generic, position 1007]

> *…You will sign in front of the incredulous eyes of your employer your declaration of nonviolent struggle. Your declaration of an all-out general strike. Full stop. Nothing else is needed. There will be 100,000 of us, and we will block Italy, offices, services, production. We will pull the plug of this infamous regime.* [Generic, position 2127]

Summary: Explaining Green Pass Opposition Without Involving Vaccines

Among those opposing the introduction of the green pass, especially among university students, only a few were in favor of vaccinations and those in favor of freedom of choice were typically hesitant about vaccines. The "being aware" antivaccine discourse has been typically dismissed by a large fraction of the Italian society and by the political class as conspiratorial in nature and not worth considering. Anti–green pass antivaccine supporters have oriented themselves toward different argumentations to defend their positions, revolving around legal aspects related to the concept of personal freedom. Our considerations are well summarized by the following message.

> *…The main argument must continue to be that one must be able to refuse an injection, whatever it may be. The body is mine and I decide. And if you were to be convinced that the serum prevents x% of the infection (as some try to suggest), would our whole battle fall apart? I certainly hope it's not the case.* [University, north, position 24367]

The battle is fought on grounds different from that of vaccines, but vaccines are what this battle is for.

Other Aspects

COVID-19

In the no-green-pass corpus, 2 main positions about COVID-19 emerged. According to the first, COVID-19 exists but is much less dangerous than what it is communicated by the mainstream media.

> *…In addition, in response to the pathetic provocation, I would like to point out that 99% of COVID deaths are of over-80s with multiple pathologies.* [University, center, position 2199-2202]

> *…COVID exists but you can't stop the world because of it. It's a fucking flu, especially for young people. Many more people have died of the flu and it has never been talked about.* [University, north, position 2864]

According to the second dominant position, COVID-19 does not exist and is yet another element of a bigger plan conceived to limit personal freedom and eradicate free thinking through fear.

> *…Do you realize that you're talking about a virus that nobody anywhere in the world can prove exists?* [University, north, position 1328]

> *…The virus has never been isolated or purified.* [University, north, position 6509]

At the junction of these 2 narratives, COVID-19 would be a strategy to pursue other means.

> *…the virus is just a means to achieve other goals that have nothing to do with health protection.* [University, center, position 8092-8095]

Expertise

If COVID-19 does not exist or is not particularly dangerous, then the need for measures such as the green pass would be unfounded. These beliefs are supported by a wide network of experts, which according to users are brave free thinkers who are not afraid of speaking their mind and standing against these covert powers.

> *…In addition, the most important doctor we have in Italy, Dr Remuzzi with H index 189, has long since drawn up an approved treatment protocol. Go to the website of the [name] Negri Institute and find out more. Dr Scoglio, candidate for the 2018 Nobel Prize, should also be considered.* [University, center, position 14640-14643]

> *…COVID can be treated at home, with medication. There is a group of volunteer doctors who do just that. "Terapie domiciliari COVID," a very popular Facebook group.* [University, south, position 1974]

> *…Listen also to Dr Citro, Dr [name] Montanari, Dr Bolgan, to what they say. [Authorities] forced people to get vaccinated with fear, and blackmailed young people with the green pass. There are many adverse reactions and they don't tell you that, so resist for your own good.* [University, center, position 4198-4200]

> *…The mask does not protect against viruses. Instead, it creates colonies of bacteria that you breathe in, as well as other filth that I won't tell you about, not to look like a conspiracy theorist. My colleague's comments on Dr Gatti are right. A great nanopathologist.* [University, north, position 742]

> *…According to Dr Delgado, it is not a virus that causes the disease. I will explain this when we meet.* [University, north, position 3485]

Preferred Measures

Among those who believe that COVID-19 is actually an issue to be contained, some try to delineate alternatives to the green pass. These include the use of masks, social distancing, tests, and dual teaching (both in presence and online).

> *…Exactly, you must respect all the rules to prevent contagion and therefore masks and distancing.* [University, south, position 1467]

> *…if we really want to be sure that the virus does not spread in the university, shouldn't the swab be used for everyone who enters the university, as it is the only instrument with a high percentage of detection of the virus?* [University, north, position 25297]

> *…However, I would like to see mixed teaching, both face-to-face and online, at least in the first semester so as not to increase the risk of infection and to allow everyone to get vaccinated. The situation in [place], with transport and everything, means that the risk of contagion is too high, even for those who are vaccinated and may be carriers. I don't feel like taking the responsibility of walking around in [place], even if I'm vaccinated, and putting other people's life at risk.* [University, center, position 2095-2102]

However, in the same groups, there is a strong critique of the dehumanization caused by online teaching, and tests are perceived as burdensome (economically and physically) and as unfair.

> *…In spite of the effort to reach the university, it is not real university what you do online. [The real one is] made of people, looks, REAL dialogues; it is precisely the effort and the time spent to go to the university that sanctions its founding and formative value. Distance learning is not an appropriate cultural medium.* [University, north, position 19204]

> *…the test becomes an economically limiting tool for the individual, since university students are not guaranteed free access to this service at all, which puts an economic burden on those who choose not to vaccinate.* [University, north, position 25298]

Antitest and Antimask Positions

Although more testing and systematic use of masks are sometimes suggested as a preferred protection strategy, many users have concerns about both. Some users think that tests and masks do not work; some think tests are dangerous as while collecting a mucus sample, it is possible to damage the brain; and some believe that masks are dangerous as they create bacterial colonies.

> *…I still don't understand... (it's rhetorical and sarcastic) why for the most contagious virus that spreads with a single droplet - with aerosol even, in the air... you have to pierce all the way to the encephalic barrier and up to the pineal gland? Maybe because otherwise you don't assimilate graphene oxide & who knows what else? Vets have long used nasal vaccination. Ps. There have been cases of rhinoliquorrhoea, ie, loss of cerebrospinal fluid, dizziness, abnormal migraines, etc, but of course, as with everything else, everything is covered up and minimized.* [University, north, position 11697-11698]

> *…The mask does not protect against viruses. Instead, it creates colonies of bacteria that you breathe in, as well as other filth that I won't tell you about, not to look like a conspiracy theorist. My colleague's comments on Dr Gatti are right. A great nanopathologist.* [University, north, position 742]

Reliance on Anecdotal Evidence

Users often bring information to support their claims. Sometimes, they provide links to blog posts but seldom to scientific studies or to statistical analyses. Sometimes, they engage with such information critically; sometimes they do not. Of note, stories based on anecdotes and personal narratives tend not to be questioned.

> *…My grandfather died with COVID. We followed what the doctors said about treatment at home for my grandmother. She survived. My grandfather wanted to follow the standard procedure instead. 2 weeks worsening. Intensive care and death.* [University, center, position 13863-13866]

> *…I spoke to a doctor from [place]. Do you know what they do to make it look like only the unvaccinated are in the ICU? When COVID patients come in, even those vaccinated with two doses, they have orders to move the vaccinated to other wards and leave the unvaccinated in the ICU.* [University, north, position 24524]

Discussion

Vaccine Skepticism and Public Health Recommendations

Our analysis clearly shows how the green pass has become a proxy and a catalyzer for vaccine skepticism. Especially during this time, people and politicians supportive of vaccines strongly oppose vaccine skepticism or denialism. The discussion about the dangers of vaccines as well as the conspiracy theories and the misinformation in general are not considered relevant and are silenced since these positions are not backed up by scientific evidence. Antivaccine supporters have come to learn that shifting their focus on the green pass allows them to bring new arguments, which are more likely to be heard, to indirectly bring arguments against the use of vaccines. In fact, questioning the validity of the green pass rather than that of vaccines is seen as less socially problematic, albeit it remains strictly connected to the discussion about vaccines. In practice, the green pass has become the fig leaf of the antivaccine movement. That said, it is also important to note that tensions and diverging narratives exist, even within the groups under analysis. As our results show, moderate positions (ie, COVID-19 is an issue, but the green pass is not an appropriate measure) coexist with conspiracy theories (ie, COVID-19 does not exist and COVID-19, vaccines, and the green pass are part of a bigger plan). De facto, opposition to the green pass is what glues together these opinions and attitudes. This opposition is often justified on the grounds of a naïve idea of freedom, conceptualized in a jurisprudential, consequentialist, or deontological form. Based on our findings, we believe that it is possible to trace some recommendations for public health

authorities and political institutions engaging with communication on these topics:

1. Acknowledge the doubts of individuals opposed to the green pass without dismissing their opinions and arguments as ramblings.
2. Disambiguate the purpose of the green pass: it should be made clear that it is a tool intended to incentivize vaccinations and thus to protect people—not only to people who cannot get vaccinated but also to protect everybody's personal freedom (ie, those who are not willing to risk to contract the virus but still desire to enjoy a meal in a restaurant, watch a theater play, or a football match in a stadium, etc). We see this discussion as a reminiscent of the long-standing debate about smoking in closed environments.
3. Since freedom is an important topic, counteract the models of freedom in which the opposition to the green pass is grounded, offering alternatives, for example, Rawls' "greatest equal liberty" principle [28], according to which each person should be given the most extensive basic freedoms that are compatible with another person's freedom.
4. Clarify the legal basis of the green pass, explaining how it is founded and regulated in existing jurisprudence, and how its scope and application is defined and limited by the contingency of the pandemic. It is necessary to explain why it has a specific "expiry date" and under which circumstances and for how long people should expect these measures to be in place.
5. Keep informing about vaccines, with a specific focus on transparency and risk-benefit balance. In this context, complement as much as possible narratives based on data and scientific evidence with personal narratives (still backed up by science), as they are easier to relate to and can be more effective [29].

Ethical Considerations and Recommendations

A Plea for Active Social Listening

Communication is a key component of human life. The ability to communicate privately with others can be understood as an expression of the right to privacy. Privacy, in turn, is not a luxury that can be easily overridden by other seemingly more urgent or more important needs. Rather, it is a fundamental human right recognized by the United Nations Declaration of Human Rights and many other international and national treaties. The COVID-19 pandemic has presented us with tricky dilemmas regarding the protection of both privacy and public health. Although there is no doubt about the need for the effective management of the pandemic, concerns have been voiced that "measures taken to control the spread of COVID-19 have negatively impacted the enjoyment of the right to privacy and other human rights" [30]. These concerns become even more acute when measures are coupled with artificial intelligence technology that can enhance not only analysis and forecasting but also the ability to effectively target the behavior of groups and individuals [31]. The key ethical question is therefore how effective communication and management during important public health crises such as pandemics is possible without undermining privacy as a human right.

Telegram grants end-to-end encryption, and encrypted communication might grant a sense of safety to users. In fact, owing to this perceived safety, often it is chosen for illegal activities as it happened for the sale of false green passes [18]. However, when a curious user acquires access to the group, either directly or with social engineering techniques, he has access to the entire history of the chat, no matter the encryption. It is worth noting that similar or related groups often cross-share messages; when a message is shared, it incorporates a link to the original chat where it was posted. Thus, by scraping chats for "t.me" links, it is rather simple to obtain access to related groups. Finally, it is important to mention that often these groups use bots offering more advanced moderation features, for example, silencing a user for a specified amount of time. As bots come to users as black boxes, it would not be difficult to load them with malicious features, for example, sending the links of the chats where they are used if specific rules fire. Even when users do not use their name and surname as their username, still there are many possible strategies for reidentification. Users might share emails, locations, and even pictures. Crossing this information and identifying a person is just a matter of amount of data, time, and commitment.

Having proven that the approach and the techniques detailed in this paper can provide useful and deep insight on critical topics debated in Telegram groups, we still tend to think that these techniques should not be applied broadly for social listening. We live in a time in which societies are already experiencing a progressive loss of trust, and techniques of passive social listening—intended as collecting information from digital communities without engaging with them—can only worsen the situation. Passive social listening, as detailed in this paper, is incredibly powerful, as it can extensively and rapidly map communities, measure their discussions, and potentially help predicting protests and violent actions. Active social listening—intended as actively asking people their opinion on delicate topics such as vaccine distribution strategies or safety measures—is slower and less comprehensive, as it depends on creating efficient bidirectional interfaces between the public and authorities. However, it has a big advantage: it can build trust rather than undermine it further. Engaging directly with communities by offering concerned people the possibility to voice their worries can create a sense of not only being listened to but also of being heard, recognized, and valued. A recent example of an active social listening tool is PubliCo [32], a web-based platform that collects data on public perception of the pandemic and its management. Following a participatory citizen science approach, it invites users not only to provide data but also to suggest new survey items or to research the database with queries of their particular interest [33].

Transparency and Recommendations

The software and the procedure we developed are subject to the dual-use problem. In nondemocratic regimes, they could be used not only to map and understand dissent but to eradicate it. Our decision to share it is motivated by 3 reasons: first, science should be open and transparent in its objectives, means, and

methods, not only in its findings. Second, as Steven Levy noted, "If you don't have access to the information you need to improve things, how can you fix them?" [34]. Pavel Durov, Telegram's founder, stated that "Telegram must continue to serve the world as an example of a tech company that strives for perfection and integrity" [35]. If Telegram wants to stay true to that claim, the company needs to know how a characteristic of their software can be exploited as a vulnerability compromising users' privacy. Third, if a nondemocratic regime would want to develop a similar system, it could do it anyways unless this vulnerability is fixed.

Limitations

As we collected our data from public Telegram groups, our sampling is not representative of the general anti–green pass population. We do not have any information about the magnitude of the phenomenon nor do we have demographic variables to stratify the analysis. However, the sample is relevant for the scope of this study and we can characterize why and how these groups oppose the green pass by drawing reliable conclusions and outlining possible approach strategies. Our approach to thematic analysis departs from standards: in thematic analysis, data should be disassembled and reassembled in a different shape to reveal its themes and patterns [36,37] with a bottom-up approach to coding. Codes should emerge during the analysis to capture emerging and unforeseen phenomena, which contrasts with the very notion of autocoding that we employed. To mitigate this, we adopted an iterative process with continuous testing, analysis, and expansion of the rules. Still, we believe autocoding is a good compromise to map the content of large volumes of data in reasonable time.

Conclusion

Through our social listening analysis on Telegram chats, we conclude that a large fraction of individuals opposed to the green pass share antivaccine views. We also show they generally do not argue their opposition to the green pass with antivaccine rhetoric but rather focus on the legal aspects and limitations of personal freedom. We suggest that public health and political institutions provide a legal explanation and a context for the use of the green pass as well as to continue focusing on vaccine communication to inform hesitant individuals. Finally, we point to the ethical ramifications of our research and propose ways to ensure that social listening analysis is transparent and ethically sound. Further work is needed to define a consensual ethical framework for social listening for public health.

Authors' Contributions

GS, FG, and NBA conceived the study design. GS wrote the software and collected the data. FG and GS performed the quantitative analysis. FG prepared the figures and GS performed the qualitative analysis. GS, FG, and NBA wrote and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Data collection: composition of the groups, users, and number of messages. The table lists the category to which the groups belong (no-green-pass or control), the description of the group, the number of users, and the number of messages in the group.
[DOCX File , 14 KB - jmir_v24i2e34385_app1.docx]

Multimedia Appendix 2
Original quotes in Italian included in the qualitative analysis.
[DOCX File , 24 KB - jmir_v24i2e34385_app2.docx]

Multimedia Appendix 3
Categories and topics included in the qualitative analysis.
[DOCX File , 21 KB - jmir_v24i2e34385_app3.docx]

Multimedia Appendix 4
The 20 most frequently used lemmas, symbols, or expressions in control chats. The table lists the most frequently used lemmas, symbols, or expressions (in percentage) on average across each individual control chat (n=5).
[DOCX File , 14 KB - jmir_v24i2e34385_app4.docx]

Multimedia Appendix 5
The 20 most used lemmas, symbols, or expressions in green pass opposition chats. The table lists the most frequently used lemmas, symbols, or expressions (in percentage) on average across each individual green opposition chat (n=5).
[DOCX File , 14 KB - jmir_v24i2e34385_app5.docx]

Multimedia Appendix 6

Average lemma frequency (in percentage) per rule in control versus green pass opposition chats. Average lemma frequency in percentage for the 10 most frequent lemmas in green pass opposition chats (black bars), compared with their relative frequency in control chats (grey bars), extracted from messages scoring a positive value for at least one of the rules.
[PNG File , 72 KB - jmir_v24i2e34385_app6.png]

References

1. Wouters OJ, Shadlen KC, Salcher-Konrad M, Pollard AJ, Larson HJ, Teerawattananon Y, et al. Challenges in ensuring global access to COVID-19 vaccines: production, affordability, allocation, and deployment. The Lancet 2021 Mar;397(10278):1023-1034. [doi: 10.1016/s0140-6736(21)00306-8]
2. Ten threats to global health in 2019. World Health Organization. 2019. URL: https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019 [accessed 2021-09-09]
3. Germani F, Biller-Andorno N. The anti-vaccination infodemic on social media: A behavioral analysis. PLoS One 2021;16(3):e0247642 [FREE Full text] [doi: 10.1371/journal.pone.0247642] [Medline: 33657152]
4. Saban M, Myers V, Ben Shetrit S, Wilf-Miron R. Issues surrounding incentives and penalties for COVID-19 vaccination: The Israeli experience. Prev Med 2021 Dec;153:106763 [FREE Full text] [doi: 10.1016/j.ypmed.2021.106763] [Medline: 34352308]
5. Wilf-Miron R, Myers V, Saban M. Incentivizing Vaccination Uptake: The "Green Pass" Proposal in Israel. JAMA 2021 Apr 20;325(15):1503-1504. [doi: 10.1001/jama.2021.4300] [Medline: 33720271]
6. EU digital COVID certificate. European Commission. 2021. URL: https://ec.europa.eu/info/live-work-travel-eu/coronavirus-response/safe-covid-19-vaccines-europeans/eu-digital-covid-certificate_en [accessed 2021-09-09]
7. DECRETO-LEGGE 23 luglio 2021, n. 105. Misure urgenti per fronteggiare l'emergenza epidemiologica da COVID-19 e per l'esercizio in sicurezza di attivita' sociali ed economiche. Gazzetta Ufficiale. 2021 Sep 18. URL: https://www.gazzettaufficiale.it/eli/id/2021/07/23/21G00117/sg [accessed 2021-09-09]
8. DECRETO-LEGGE 22 aprile 2021, n. 52. Misure urgenti per la graduale ripresa delle attivita' economiche e sociali nel rispetto delle esigenze di contenimento della diffusione dell'epidemia da COVID-19. Gazzetta Ufficiale. 2021 Apr 23. URL: https://www.gazzettaufficiale.it/eli/id/2021/04/22/21G00064/sg [accessed 2021-09-09]
9. Kirby P. COVID passes set to stay as Europe heads for winter. BBC News. URL: https://www.bbc.com/news/world-europe-58973334 [accessed 2021-10-28]
10. Kang-Xing J. Keeping people safe and informed about the coronavirus. Meta. 2020. URL: https://about.fb.com/news/2020/12/coronavirus/ [accessed 2021-09-09]
11. Coronavirus: Staying safe and informed on Twitter. Twitter Blog. 2021. URL: https://blog.twitter.com/en_us/topics/company/2020/covid-19 [accessed 2021-09-09]
12. Rogers R. Deplatforming: Following extreme Internet celebrities to Telegram and alternative social media. European Journal of Communication 2020 May 06;35(3):213-229. [doi: 10.1177/0267323120922066]
13. Jakobsen J. A practical cryptanalysis of the Telegram messaging protocol. Aarhus University. 2015. URL: https://enos.itcollege.ee/~edmund/materials/Telegram/A-practical-cryptanalysis-of-the-Telegram-messaging-protocol_master-thesis.pdf [accessed 2021-09-10]
14. Telegram FAQ. Telegram. 2021. URL: https://telegram.org/faq?setln=en [accessed 2021-09-10]
15. Urman A, Katz S. What they do in the shadows: examining the far-right networks on Telegram. Information, Communication & Society 2020 Aug 20:1-20. [doi: 10.1080/1369118x.2020.1803946]
16. Dargahi Nobari A, Sarraf MHKM, Neshati M, Erfanian Daneshvar F. Characteristics of viral messages on Telegram; The world's largest hybrid public and private messenger. Expert Systems with Applications 2021 Apr;168:114303. [doi: 10.1016/j.eswa.2020.114303]
17. Shapiro J. Managing violent covert organizations. In: The Terrorist's Dilemma. Princeton: Princeton University Press; 2013.
18. Falsi green pass in vendita, bloccati 32 canali telegram. AGI. 2021. URL: https://www.agi.it/cronaca/news/2021-08-09/falsi-green-pass-su-telegram-blitz-polizia-13528269/ [accessed 2021-09-10]
19. Rossi A. Il movimento nato su Telegram per dire no al green pass. AGI. 2021. URL: https://www.agi.it/cronaca/news/2021-08-31/green-pass-movimento-telegram-13721900/ [accessed 2021-09-10]
20. Federal act on research involving human beings (human research act, HRA). The Swiss Federal Council. 2014. URL: https://www.admin.ch/opc/en/classified-compilation/20061313/index.html [accessed 2021-09-10]
21. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance). EUR-Lex. 2016. URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj [accessed 2021-09-10]
22. de Sterlich P. Lista nomi italiani. GitHub Gist. 2017. URL: https://gist.github.com/pdesterlich/2562329 [accessed 2021-09-10]
23. Codici statistici delle unità amministrative territoriali: comuni, città metropolitane, province e regioni. ISTAT. 2021. URL: https://www.istat.it/it/archivio/6789 [accessed 2021-09-10]
24. spaCy 101: everything you need to know. spaCy Usage Documentation. 2021. URL: https://spacy.io/usage/spacy-101 [accessed 2021-09-21]

25. Liu B. Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies 2012 May 23;5(1):1-167. [doi: 10.2200/s00416ed1v01y201204hlt016]
26. Bianchi F, Nozza D, Hovy D. FEEL-IT: emotion and sentiment classification for the Italian language. 2021 Presented at: Eleventh Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis; April 19; Online.
27. Spitale G, Germani F, Biller-Andorno N. The TSL machine: parser, lemma analysis, sentiment analysis and autocoding for Telegram chats. Zenodo. 2021. URL: https://zenodo.org/record/5534045 [accessed 2021-09-29]
28. Rawls J. A Theory of Justice: Original Edition. Harvard: Harvard University Press; 2005.
29. Drewniak D, Glässel A, Hodel M, Biller-Andorno N. Risks and Benefits of Web-Based Patient Narratives: Systematic Review. J Med Internet Res 2020 Mar 26;22(3):e15772 [FREE Full text] [doi: 10.2196/15772] [Medline: 32213468]
30. Cannataci J. Report of the special rapporteur on the right to privacy (A/76/220). United Nations General Assembly. 2021. URL: https://undocs.org/A/76/220 [accessed 2021-10-20]
31. The right to privacy in the digital age: report 2021. Office of the United Nations High Commissioner for Human Rights. 2021. URL: https://www.ohchr.org/EN/Issues/DigitalAge/Pages/cfi-digital-age.aspx [accessed 2021-10-20]
32. Spitale G, Merten S, Jafflin K, Schwind B, Kaiser-Grolimund A, Biller-Andorno N. PubliCo. 2020. URL: https://publico.community [accessed 2021-09-09]
33. Spitale G, Merten S, Jafflin K, Schwind B, Kaiser-Grolimund A, Biller-Andorno N. A Novel Risk and Crisis Communication Platform to Bridge the Gap Between Policy Makers and the Public in the Context of the COVID-19 Crisis (PubliCo): Protocol for a Mixed Methods Study. JMIR Res Protoc 2021 Nov 01;10(11):e33653 [FREE Full text] [doi: 10.2196/33653] [Medline: 34612823]
34. Levy S. Hackers: Heroes of the Computer Revolution. 25th Anniversary Edition. Sebastopol: O'Reilly Media; 2010.
35. Durov P. Message 142. Telegram: Durov's Channel. 2020. URL: https://t.me/durov/142 [accessed 2021-09-29]
36. Braun V, Clarke V, Hayfield N, Terry G. Thematic analysis. In: Handbook of Research Methods in Health Social Sciences. Singapore: Springer; 2018.
37. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]

Abbreviations

GDPR: General Data Protection Regulation

Edited by M Focsa; submitted 20.10.21; peer-reviewed by E Furlan, H Ayatollahi; comments to author 28.11.21; revised version received 02.12.21; accepted 09.12.21; published 16.02.22.

Please cite as:
Spitale G, Biller-Andorno N, Germani F
Concerns Around Opposition to the Green Pass in Italy: Social Listening Analysis by Using a Mixed Methods Approach
J Med Internet Res 2022;24(2):e34385
URL: https://www.jmir.org/2022/2/e34385
doi:10.2196/34385
PMID:35156930

Original Paper

Patients' Perspectives on Transforming Clinical Trial Participation: Large Online Vignette-based Survey

Van Thu Nguyen[1,2], MPH, DPhil; Philippe Ravaud[1,3,4], MD, DPhil; Viet Thi Tran[1,4], MD, DPhil; Bridget Young[2], DPhil; Isabelle Boutron[1,3,4], MD, DPhil

[1]Université de Paris, Centre of Research Epidemiology and Statistics, Inserm, Paris, France

[2]Department of Public Health, Policy and Systems, Institute of Population Health, University of Liverpool, Liverpool, United Kingdom

[3]Cochrane France, Paris, France

[4]Centre d'Epidémiologie Clinique, Hôpital Hôtel Dieu, Assistance Publique des Hôpitaux de Paris, Paris, France

Corresponding Author:
Van Thu Nguyen, MPH, DPhil
Université de Paris
Centre of Research Epidemiology and Statistics
Inserm
1 Parvis de Notre Dame
Paris, 75004
France
Phone: 33 605714478
Email: nguyenthuvandkh@gmail.com

Abstract

Background: Patients' participation is crucial to the success of randomized controlled trials (RCTs). However, recruiting and retaining patients in trials remain a challenge.

Objective: This study aims to describe patients' preferences for the organization of RCTs (visits on- site or remotely) and evaluate the potential impact of fulfilling preferences on their willingness to participate in a clinical trial.

Methods: This was a vignette-based survey. Vignettes were case scenarios of real clinical trials assessing pharmacological treatments. These RCTs evaluated 6 prevalent chronic diseases (ie, osteoporosis, osteoarthritis, asthma, cardiovascular diseases, diabetes, and endometriosis). Each vignette described (1) the RCT and characteristics of the treatment tested (ie, doses, administration routes) and (2) the trial procedures and different options (on-site or remotely) for how the trial was organized for informed consent, follow-up visits, and communication of results when the trial was completed. We recruited 628 participants from ComPaRe (www.compare.aphp.fr), a French e-cohort of patients with chronic diseases. The outcomes were the participants' preferences for the way the trial was organized (on-site or remotely) and their willingness to participate in the trial.

Results: Of the 628 participants who answered the vignettes, 491 (78.2%) were female (median age 55 years), with different chronic diseases ranging from endometriosis in 59 of 491 (12%) patients to asthma in 133 of 628 (21.2%) patients. In addition, 38 (6.1%) participants wanted to provide informed consent and all trial visits on-site, 176 (28%) wished to participate in the trial entirely remotely, and 414 (65.9%) wanted to combine remote-based and hospital-based visits. Considering the trial as a whole, when the trial was organized in a way that the patients preferred, the median (Q1-Q3) likelihood of participation in the trial was 90% (80-100) versus 60% (30-80) if the trial followed the patients' nonpreferred model. Furthermore, 256 (40.8%) patients responded to open-ended questions expressing their experience with trial participation and visits to the hospital and providing suggestions for improvement. The patients emphasized the need to personalize the way a trial is organized according to each patient's needs and conditions.

Conclusions: There was a significant diversity in the participants' preferences. Most participants preferred hybrid organization involving both on-site and remote visits. Participants were more likely to participate in a trial organized according to their preferences.

(J Med Internet Res 2022;24(2):e29691) doi:10.2196/29691

KEYWORDS

randomized controlled trial; remote trial; telemedicine; patient experience; trial participation; RCTs; participation; recruitment; patient preferences; remote medicine; pharmacological treatments

Introduction

Patients' participation is crucial to the success of randomized controlled trials (RCTs). However, recruiting and retaining patients in trials remain a challenge [1,2]. For example, less than a third of RCTs funded by the National Institute for Health Research (NIHR) in the United Kingdom achieved the target sample size, and about a third of RCTs in the United States recruited less than 75% of the planned sample size [1,3]. With the increasing complexity of trial procedures, patients' decisions to take part and remain in a trial depend not only on the potential benefits and risks of the interventions evaluated but also on the practical logistics of trials and how burdensome these are to patients [4,5]. Evidence shows that participating in RCTs can be burdensome for patients, particularly due to the ways informed consent is managed, follow-up visits organized, and trial results communicated [5,6]. Several efforts have tried to ease this burden by using technologies to allow patients to take part in a trial remotely using electronic informed consent, dispensing study drugs at patients' homes, performing entirely web-based data collection, and conducting virtual visits via video calls [7-9]. However, the model of remote trials has not significantly succeeded in improving trial recruitment and retention and might not be suitable to all patient groups [7,10].

Information about patients' preferences for the way a trial is organized could inform the trial planning to enhance the participation rate. However, the previous literature on patients' willingness to participate in clinical trials mainly focused on their attitudes toward the randomization, without considering other aspects of trial organization [11]. In this study, we developed a new approach, using an online vignette-based survey to involve a large number of patients to explore their preferences on trials visits (on-site or remotely) and to evaluate the impact of fulfilling their preferences on potential participation.

Methods

Study Design

We conducted a vignette-based survey asking patients a set of directed questions to elicit their preferences for different ways of organizing a trial. Vignettes have traditionally been used in a number of areas, including medical training to evaluate clinical practice, and are being increasingly used in research to address topics such as identifying the best trial designs for methodological questions [12-15]. In this study, vignettes were case scenarios of real clinical trials assessing pharmacological treatments. These vignettes explained to patients what a clinical trial is and what they are expected to do when participating in the clinical trial.

This study received ethical approval from the French National Institute of Health and Medical Research Ethic Committee (IRB00003888; reference 19-580). We reported our findings following the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) checklist for online surveys [16].

Vignette Development

The vignettes were developed from existing protocols of ongoing or recently completed RCTs either published or available on clinicaltrials.gov. Two patient representatives and a steering committee comprising trialists and methodologists contributed to the development of the vignettes.

Clinical Trial Protocol Search

We systematically searched clinicaltrials.gov and PubMed for protocols of ongoing or recently completed (2017 onward) phase 3 RCTs evaluating pharmacological treatments. We limited the search to 6 prevalent chronic diseases: osteoporosis, osteoarthritis, asthma, cardiovascular diseases, diabetes, and endometriosis (Multimedia Appendix 1). Trials were included if they had a parallel design, at least 1 year of follow-up, and a full protocol available. We excluded trials that were exclusively on patients less than 18 years old or that were conducted exclusively in Asia, Africa, and Latin America. We also excluded trials testing treatments for secondary conditions (eg, osteoporosis induced by glucocorticoids). In total, 93 RCTs were retrieved from the search, with 18 (19%) fulfilling the inclusion criteria. One protocol was chosen for each disease, with the goal of maximizing the diversity of funding and the type of administration route across the vignettes, leading to 6 protocols being chosen for vignette development (Multimedia Appendix 2).

Vignette Conception

Each vignette was structured in 2 parts. The first part described the RCT and the characteristics of the treatment tested (ie, doses, administration routes). The second part described the procedures of the trial and different options of how the trial was organized for each step: (1) informed consent, (2) follow-up visits, and (3) communication of results when the trial was completed. For informed consent, the participants were asked to choose 1 of the 3 ways: (1) The trial was explained to the patients, and they signed the informed consent form at the research center; (2) information was explained via a video, and the patients electronically provided signed informed consent; and (3) the trial was explained to the patients at the research center, and they then electronically signed the informed consent form at home after getting an opportunity to discuss with their family. For follow-up visits, we provided details of the number of visits and types of tests involved in each visit. The patients were asked to choose 1 of 3 options: (1) All follow-up visits happened at the research center; (2) all follow-up visits happened at the patients' homes, with a nurse visiting their home, and tests were conducted at a local hospital if they could not be performed at home; and (3) the patients were able to decide which visits happened at the research center and which at home. We described all logistics involved in each option (eg, contact with doctor, travel and waiting time). For communication of results,

the patients were asked to choose 1 of 4 options: (1) have a face-to-face meeting with investigators, (2) have a video call with investigators, (3) receive the results by email, and (4) receive the results by post. The description of each option for informed consent and communication of results were kept unchanged across the vignettes, while the number of visits and frequency of tests performed varied according to the protocols of the real RCTs. Textbox 1 presents an example of the choices for the informed consent process for the osteoarthritis RCT. Other vignettes are provided in Multimedia Appendix 3. The vignettes were then translated into French by a professional translator. A senior researcher (author IB) and a patient representative who was a native French speaker reviewed the translation to ensure its accuracy.

Textbox 1. Vignette for the randomized controlled trial (RCT) of osteoarthritis.

1. Informed consent

Before you participate in the clinical trial, the research team will explain to you the study, the new treatment, and the schedule of assessments. You will sign a consent form if you agree to participate.

Where do you want to have the information of the clinical trial explained to you and sign the consent form? Choose 1 of the 3 choices below:

At the university hospital

You will:

- Travel to the hospital and wait to see a study doctor.
- Meet the doctor who explains the trial to you.
- Ask your questions to the doctor.
- Sign the consent form if you agree to participate.
- Keep a copy of the consent form.

At home using the internet

You will:

- Stay at home.
- Watch a video online explaining the trial.
- Call the study doctor by telephone if you have questions when you want during working hours.
- Discuss with your family.
- Sign the consent form via a computer when you are ready.

A copy of the consent form will be sent to you by email or by post based on your choice.

At the university hospital and at home

You will:

- Travel to the hospital and wait to see a doctor.
- Meet the doctor who explains the trial to you.
- Ask your questions to the doctor.
- Return home and discuss with your family.
- Call the study doctor by telephone if you have questions when you want during working hours.
- Sign the consent form via a computer when you are ready.

A copy of the consent form will be sent to you by email or by post based on your choice.

Participants

We recruited patients from La communauté de patients pour la recherche (ComPaRe, an ongoing cohort of patients with chronic conditions in France) [17]. The patients were adults (>18 years old) who reported having at least 1 chronic condition (defined as a condition requiring health care for at least 6 months). They helped accelerating research on their conditions by completing patient-reported outcome questionnaires, suggesting ideas for new research or participating in the analysis of research projects [18-20]. By February 2020, there were 36,000 patients participating in ComPaRe. All patients provided electronic informed consent before participating in the e-cohort. By signing this electronic informed consent form to participate in ComPaRe, they agreed to receive invitations to participate in research provided via the platform.

We sent an invitation email to all patients who reported having asthma, diabetes, endometriosis, hypercholesterolemia, osteoarthritis, or osteoporosis via the ComPaRe platform. Two reminder messages were sent to nonrespondents. We then sent

the questionnaire including the vignette corresponding to the patients' chronic diseases to patients who replied yes to the invitation email via the secure ComPaRe platform. The patients were recruited from February 12 to April 23, 2020.

Outcomes

We evaluated the following outcomes:

- Participants' preferences for the way a trial is organized at each step: The patients were asked to indicate their preferred choices for the following questions: *Where would you like to receive the information about the trial and sign the informed consent? Where would you like to do the follow-up visits? How do you want to receive the results of the trial?*
- Participants' willingness to participate in the trial: The patients were asked the following questions: *If this step is performed at the hospital, how likely would you participate in the trial? If this step is performed at your home, how likely would you participate in the trial? If this step is performed at both the hospital and your home, how likely would you participate in the trial?* The patients answered each question on a scale from 0 to 100, with 0 being very unlikely and 100 being very likely.
- The patients were also invited to propose new ideas for organizing each step of the trial.

They were able to express their ideas in free text by responding to open-ended statements: *If you have an idea to improve your experience with receiving information about the trial and sign informed consent, please let us know* and *If you have an idea to improve your experience with follow-up visits, please let us know*.

Data Analysis

Descriptive analysis was used to describe the characteristics of the patients participating in the survey and their preferences for trial organization. The analysis was performed using R software (version 4.0.2).

Analysis of the patients' answers to open-ended questions about their ideas for how clinical trials should be organized were informed by thematic analysis. Data were imported into NVivo (QSR International) to facilitate the coding process. One researcher (author VN) performed open coding through an iterative process. A subset of responses to open-ended questions from 50 participants was coded line by line. The codes were then organized into themes to create a coding framework. This coding framework was then discussed with a senior researcher (author IB) and refined by considering the context of the data with the wider data set. The coding framework was used to analyze the remaining survey responses and refined throughout the process of analysis.

Data Sharing

De-identified data from this study are available on request to the academic researchers who have submitted a protocol to the scientific committee of ComPaRe and signed a data use agreement.

Patient and Public Involvement

Two patient representatives participated in the development of the vignettes.

Results

Study Population

We sent invitation emails to 2155 patients in the ComPaRe e-cohort, including 434 (20.14%) patients with asthma, 317 (14.71%) patients with diabetes, 480 (22.27%) patients with endometriosis, 114 (5.29%) patients with hypercholesterolemia, 276 (12.81%) patients with osteoarthritis, and 534 (24.78%) patients with osteoporosis. Of the 2155 patients, 834 (38.7%) initially responded positively to our invitations. We then sent this group the survey containing the vignettes of trials corresponding to their conditions. A total of 628 (75.3%) patients answered the vignette-based survey. Table 1 presents the demographic information of the respondents. The patients mainly lived in France (621/628, 98.9%), with age ranging from 21 to 84 years (median 55 years, IQR 44-64), and were mainly female (491, 78.2%). In addition, 427 of 628 (68%) patients lived in an urban area. Most of the patients had obtained at least a high school diploma, and 107 (17%) had experience of participating in an RCT before. Nearly 377 (60%) of the patients could reach a university hospital within 1 hour of driving from their place of residence. Multimedia Appendix 4 provides the demographic information of nonrespondents.

Table 1. Characteristics of patients.

Characteristic	Total (N=628)	Asthma (n=133)	Diabetes (n=83)	Endometriosis (n=59)	Hypercholesterolemia (n=76)	Osteoarthritis (n=125)	Osteoporosis (n=152)
Gender, n (%)							
Female	491 (78)	107 (80)	41 (49)	59 (100)	35 (46)	97 (78)	152 (100)
Male	137 (22)	26 (20)	42 (51)	0	41 (54)	28 (22)	0
Age (years), median (IQR); min-max	55 (IQR 44-64); 21-84	45 (IQR 36-52); 22-84	54 (IQR 46-63); 26-81	38 (IQR 32-45); 21-60	61 (IQR 56-69); 25-80	57 (IQR 50-66); 26-80	60 (IQR: 55-64); 23-83
Employment[a], n (%)							
Unemployed	51 (8.1)	18 (13.5)	2 (2)	7 (12)	4 (5)	10 (8)	10 (6.6)
Apprentice	21 (3.3)	4 (3)	3 (4)	9 (15)	0	3 (2.4)	2 (1.3)
Employed	272 (43.3)	71 (53.4)	43 (52)	39 (66)	25 (33)	41 (32.8)	53 (34.9)
Retired	169 (26.9)	15 (11.3)	24 (29)	0	37 (49)	41 (32.8)	52 (34.2)
Disabled	102 (16.2)	19 (14.3)	10 (12)	4 (7)	10 (13)	28 (22.4)	31 (20.4)
Other	12 (1.9)	5 (3.8)	1 (1)	0	0	2 (1.6)	4 (2.6)
Highest level of education[a], n (%)							
No formal diploma	14 (2.2)	4 (3)	3 (4)	1 (2)	3 (4)	2 (1.6)	1 (0.7)
Highschool diploma	99 (15.7)	18 (13.5)	18 (22)	6 (10)	13 (17)	24 (19.2)	20 (13.2)
Vocational training	76 (12.1)	17 (12.8)	11 (13)	4 (6.8)	11 (15)	11 (8.8)	22 (14.5)
Undergraduate/post-graduate	435 (69.2)	93 (69.9)	49 (59)	48 (81)	49 (64)	88 (70.4)	108 (71.1)
Other diplomas	4 (0.6)	1 (0.7)	2 (2)	0	0	0	1 (0.7)
Living area							
Rural area	201 (32)	45 (33.8)	33 (40)	20 (34)	20 (26)	40 (32)	43 (28.3)
Urban area	427 (68)	88 (66.2)	50 (60)	39 (66)	56 (74)	85 (68)	109 (71.7)
Distance to a university hospital, n (%)							
<1 hour	362 (57.6)	74 (55.6)	53 (64)	35 (59)	45 (59)	80 (64)	75 (49.3)
1-2 hours	229 (36.4)	51 (38.3)	27 (33)	20 (34)	26 (34)	37 (29.6)	68 (44.7)
2-5 hours	37 (5.9)	8 (6)	3 (3)	4 (7)	5 (7)	8 (6.4)	9 (5.9)
Previous participation in a trial, n (%)							
Yes	106 (16.9)	26 (19.5)	16 (19)	4 (7)	14 (18)	19 (15.2)	27 (17.7)
No	522 (83.1)	107 (80.5)	67 (81)	55 (93)	62 (82)	106 (84.8)	125 (82.3)

[a]One participant did not answer this question.

Patients' Preferences for the Way a Trial Is Organized

All patients answered the vignette of the trial testing treatment for their conditions. For the informed consent process, 311 (49.5%) patients indicated that they preferred to be given information about the trials and sign the consent form at home over the internet, while 239 (38.1%) patients preferred the 2-step approach: having information about the trial explained at the hospital and signing the consent form at home.

Regarding follow-up visits, 251 (39.9%) patients wished to have all follow-up visits at home and 254 (40.4%) patients preferred the combination of both on-site visits at research centers and home-based visits with the possibility to arrange the visits according to their choices. Only 122 (19.4%) patients chose to have all follow-up visits at the hospital.

Most of the respondents (276/628, 43.9%) wished to have an in-person meeting with research investigators when receiving the results of the trials, 192 (30.6%) chose to receive a summary of results by email, and 126 (20.1%) preferred a video call with research investigators (Table 2). Overall preferences were consistent across conditions. However, for the informed consent process, 33 of 59 (56%) patients with endometriosis preferred the trial to be explained to them at the hospital and to sign the informed consent form at home. Patients with hypercholesterolemia were the only group in which most patients chose to receive trial results by mail (33/76, 43%), while the other groups wished to meet a research investigator in person (Table 2).

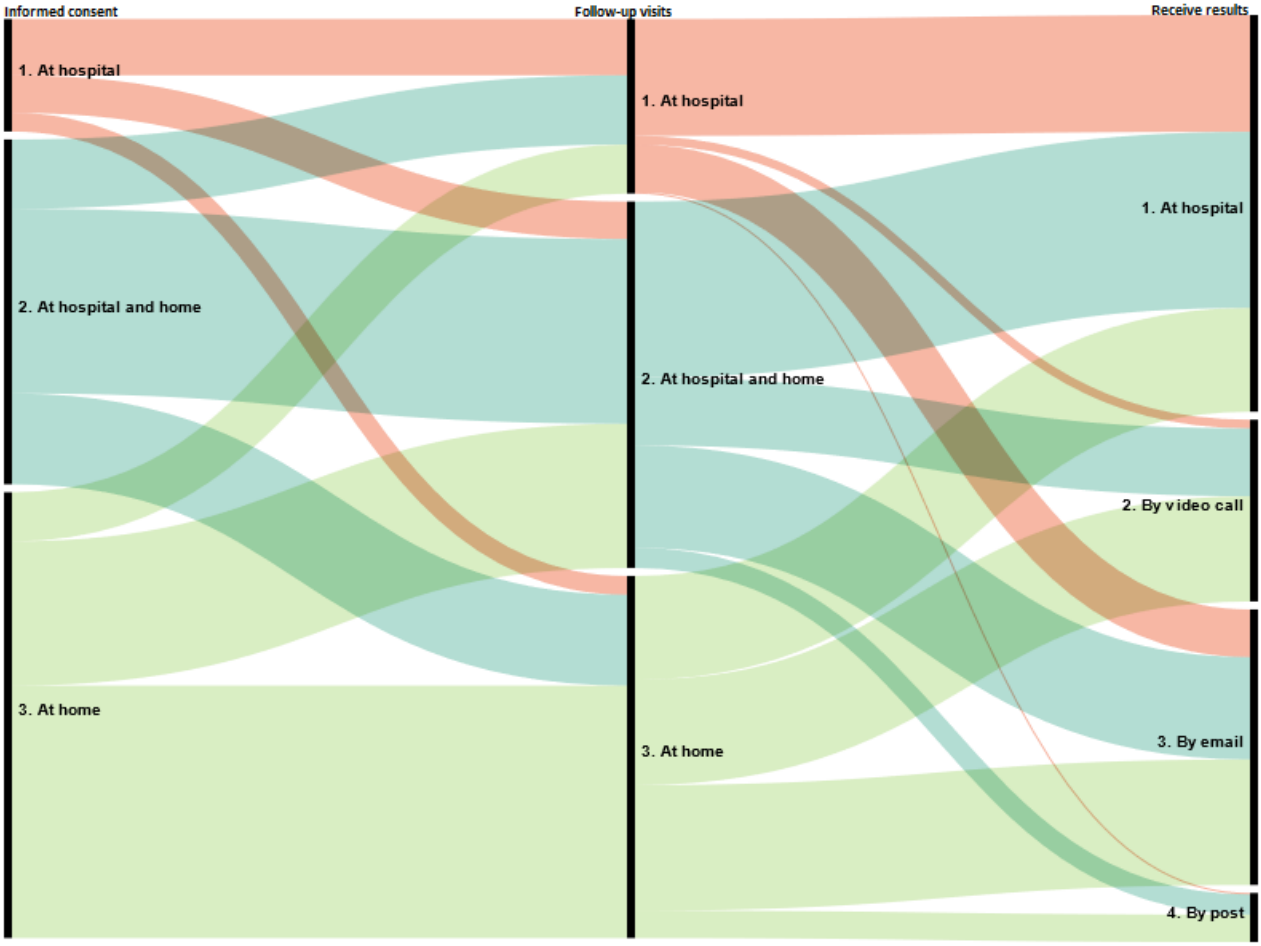
Table 2. Patients' choices of trial organization.

Choice	Total (N=628)	Asthma (n=133)	Diabetes (n=83)	Endometriosis (n=59)	Hypercholesterolemia (n=76)	Osteoarthritis (n=125)	Osteoporosis (n=152)
Informed consent, n (%)							
At home	311 (49.5)	73 (54.9)	39 (47)	19 (32)	47 (62)	58 (46.4)	75 (49.3)
At hospital and at home	239 (38.1)	32 (24.1)	40 (48)	33 (56)	22 (29)	40 (32)	62 (40.8)
At hospital	78 (12.4)	18 (13.5)	4 (5)	7 (12)	7 (9)	27 (21.6)	15 (9.9)
Follow-up visits[a], n (%)							
By choices	254 (40.4)	51 (38.3)	29 (35)	28 (48)	23 (30)	61 (48.8)	62 (40.8)
At home	251 (39.9)	58 (43.6)	42 (51)	19 (32)	41 (54)	30 (24)	61 (40.1)
At hospital	122 (19.4)	23 (17.3)	12 (15)	12 (20)	12 (16)	34 (27.2)	29 (19.1)
Receive results, n (%)							
Meeting a doctor at the hospital	275 (43.8)	58 (43.6)	39 (47)	32 (54)	24 (32)	62 (49.6)	60 (39.5)
Video call with a doctor	126 (20.1)	30 (22.6)	19 (23)	16 (27)	16 (21)	19 (15.2)	26 (17.1)
By mail	192 (30.6)	38 (28.6)	25 (30)	9 (15)	33 (43)	38 (30.4)	49 (32.2)
By post	34 (5.4)	7 (5.3)	0	2 (3)	3 (4)	6 (4.8)	16 (10.5)

[a]One participant did not answer this question.

Figure 1 illustrates the diversity of the patients' preferences regarding trial participation as a whole. For example, of 311 (49.5%) patients who wished to have the informed consent process at home, more than half (175/311, 56.3%) chose to have all follow-up visits at home. Of these 175, 73 (41.7%) wished to receive trial results by mail, 55 (31.4%) wished to be informed about trial results via a video call, and 31 (17.7%) preferred to meet a research investigator in person.

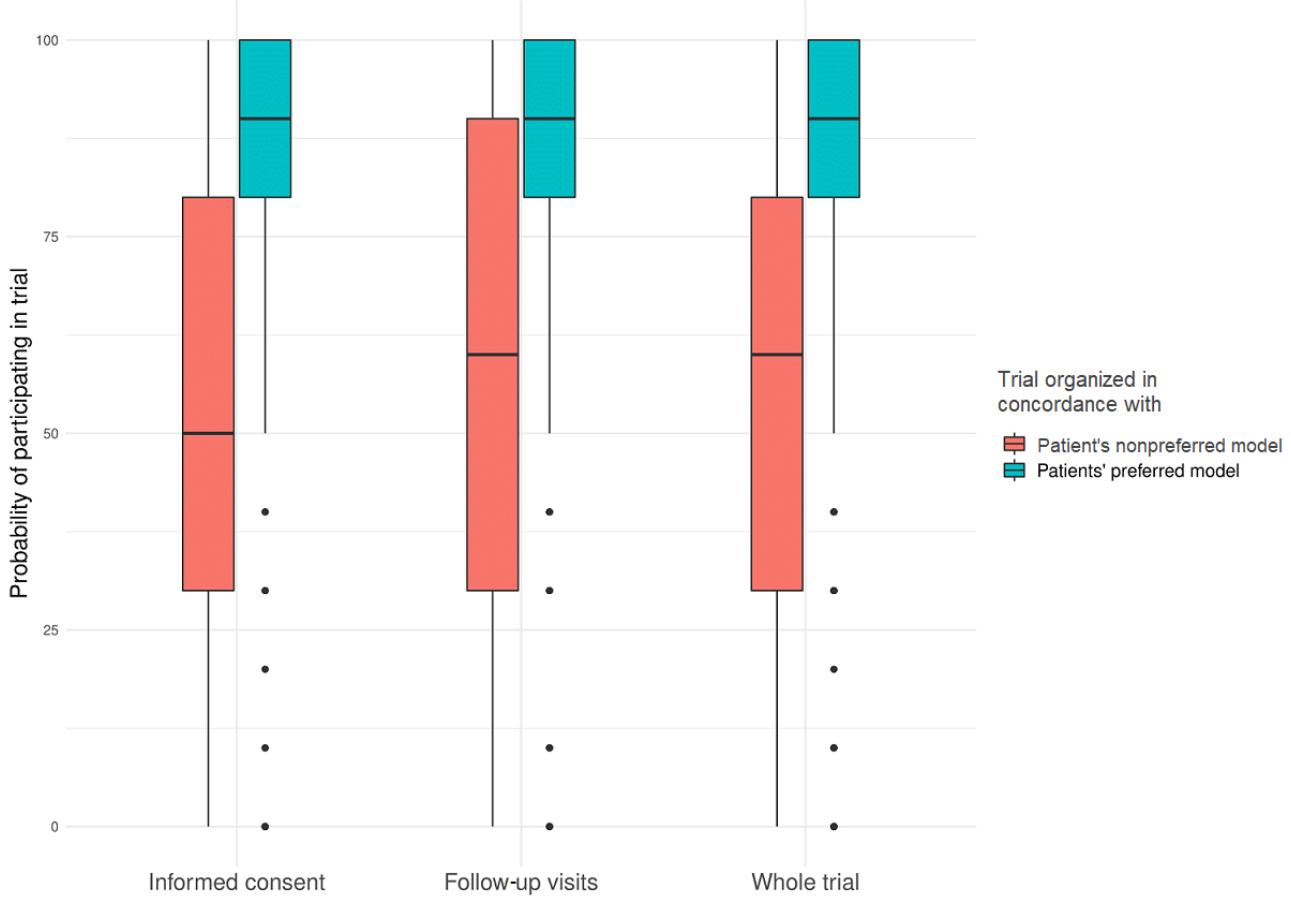
Figure 1. Diversity of patients' preferences for the way a trial is organized. This alluvial diagram presents patients' choices for each step of trial participation. The streams connecting the columns present the proportion of patients selecting each option in each column. For example, the red streams connecting "Informed consent" and "Follow-up" present the proportion of patients deciding to have follow-up visits at the hospital (n=39, 50%), at both the hospital and home (n=26, 33%), or at home (n=13, 17%) among all patients who decided to provide informed consent at the hospital (n=78). The red streams connecting "Follow-up" and "Results communication" present the proportion of patients deciding to receive the results at the hospital (n=81, 66.9%), by video call (n=6, 4.9%), by email (n=33, 27.3%), and by post (n=1, 0.8%) among all patients who decided to have all follow-up visits at the hospital (n=121).



Patients' Willingness to Participate in a Trial According to Their Preferences for Trial Organization

When the informed consent process was organized in a way that the patients preferred, the median likelihood to participate in the trial was 90% (IQR 80%-100%) versus 50% (IQR 30%-80%) if the informed consent process followed the patients' nonpreferred model. Similarly, the median likelihood to participate in a trial if follow-up visits followed the patients' preferred model was 90% (IQR 80%-100%) versus 60% (IQR 30%-90%) if trials were set up according to the patients' nonpreferred model. Considering the trial as a whole, when the trial was organized in a way that the patients preferred, the median likelihood of participation in the trial was 90% (IQR 80%-100%) versus 60% (IQR 30%-80%) if the trial followed the patients' nonpreferred model. See Figure 2.

Figure 2. Probability of participating when a trial is performed in accordance with patients' preferences.



Patients' Suggestions to Improve the Way a Trial Is Organized

Of 628 patients, 256 (40.8%) responded to at least 1 open-ended question, expressing their experience with trial participation and visits at the hospital and providing suggestions for improvement. Patients emphasized the need to personalize the way a trial is organized according to each patient's needs and conditions.

> *I think you cannot generalize, but for each clinical trial, the patient must be given a choice of how to participate. This depends mainly on the distance between home and hospital and of course whether the person has a professional activity or not. The way of participating could be proposed to the patient at the same time as the consent and the patient will then be in control of whether he or she can and wants to participate.* [Patient #125 with asthma]

Patients' Propositions to Improve Trial Visits at the Hospital

The patients suggested changes to the logistical organization of hospital visits to reduce waiting time and also suggested that flexible appointment times be offered to suit the individual situations of patients (Textbox 2).

Textbox 2. Patients' proposals to improve visits.

1. Keep appointment times and reduce waiting time.

> *Make sure that appointments with doctors or ECG radiography departments are on time.* [Patient #41 with osteoarthritis]
>
> *It all depends on the location of the hospital and how easy it is to get there by public transport (I don't drive). On the other hand, please respect the appointment times very strictly.* [Patient #5 with osteoporosis preferring home-based visits]

2. Arrange a reception dedicated to trial participants.

> *Make sure that appointments with doctors or ECG radiography departments are on time, without going through the general reception of the hospital . . . In order for me to participate in a study, the "logistics" must be as fluid as*

possible and outside the traditional care circuit in terms of administration and waiting time. [Patient #41 with osteoarthritis preferring combination of hospital and home-based visits]

3. Provide flexibility of appointment time.

Having the possibility to have intelligent appointments, to have all examination and tests in the morning or in the afternoon or from 10:00 to 15:00, for example . . . this allows fragile, sick, and tired people to take time and take care of their health, when they come from far away or when they have difficulty to move. It is important to be able to organize according to our conditions. [Patient #13 with diabetes preferring combination of hospital and home-based visits]

I would be willing to go to the hospital without any worries, but I do not want this to be done during my working hours as it should not be the concern of my employer. [Patient #65 with asthma preferring combination of hospital and home-based visits]

4. Combine follow-up visits with routine care visits.

Should we combine the visit with the examination and radiography for osteoarthritis? [Patient #97 with osteoarthritis preferring combination of hospital and home-based visits]

5. Reimburse transportation fees and provide free parking.

The fee of transportation and parking should be reimbursed for traveling to and parking at the research center. [Patient #35 with osteoarthritis preferring combination of hospital and home-based visits]

Patients' Proposals Regarding Remote Trial Visits

The patients highlighted the advantages of participating in a trial from home over the internet as a solution to reduce travel, save time, and avoid disruption of their work. They, nevertheless, indicated that human contact is important and an in-person conversation with research investigators would help reassure them when making decisions related to trial participation.

In the context of a clinical trial, a contact with a real human is important. The internet does not transmit the emotion. [. . .] It is reassuring [in] the hospital setting. They (doctors) can see my condition, and I also feel that I am a stakeholder and an actor of my own decisions when having a human in front of me. [Patient #73 with osteoarthritis preferring hospital-based visits]

One patient also explained that she wanted to keep her home as a private place for rest and recreation.

My home is a place of conviviality, rest, or recreation. I do not want my home to become a place of care. I already have auto-injections. I prefer to go to the doctor, in a center of care, even if that seems more constraining. [Patient #4 with asthma preferring combination of hospital and home-based visits]

Several patients suggested that researchers should foresee measurements to ensure the accuracy of tests and data collected outside the context of the hospital, which might influence the quality of research. Further, they also reminded researchers about the fact that not all patients would have access to equipment for video calls with doctors and that their internet connection might be unstable.

The patients also suggested that trial investigators could collaborate with local hospitals and laboratories to organize visits and examinations, which would reduce travel time, while maintaining the quality of data collected. The patients also spoke of the possibility to involve their primary care doctors in the trials, which would help them feel more reassured when participating in trials (Textbox 3).

Textbox 3. Patients' suggestions regarding remote trial visits.

1. Involve local hospitals and health care providers for follow-up visits.

I participated in a clinical trial. The appointments with the doctor took place at the hospital. The biological tests between appointments at the hospital were done at a laboratory near my home. I appreciated this organization. [Patient #56 with asthma preferring hospital-based visits]

To not wait too long at the hospital, and to be able to do the visits at a hospital nearby to reduce the travel time. [Patient #123 with asthma preferring combination of hospital and home-based visits]

2. Involve primary care doctors for informed consent and follow-up visits.

Another suggestion is to involve the primary care doctor as an intermediary to explain the study. [Patient #82 with osteoarthritis preferring hospital-based visits]

To involve the primary care doctor to avoid a part of the travel to the hospital. [Patient #51 with endometriosis preferring hospital-based visits]

Follow-up of the trial by primary care doctor and nurse for usual blood examination in close contact with the research team of the university hospital. [Patient #37 with hypercholesterolemia preferring home-based visits]

3. Ensure close contact with investigators.

Contact by video call rather than by telephone. [Patient #55 with diabetes preferring home-based visits]

It would be good to be able to exchange email with the doctor. [Patient #63 with diabetes preferring combination of hospital and home-based visits]

4. Provide equipment suitable to patients' conditions.

I have a computer with a large screen, but my osteoarthritis makes me suffer so much that I cannot sit upright for more than 5 minutes. I can use [the] telephone for a video call, and I can lie down, although it will be less effective for communication. [Patient #54 with osteoarthritis preferring hospital-based visits]

I would like to have home-based visits, but I am not sure to have a webcam with my computer. [Patient #6 with hypercholesterolemia preferring combination of hospital and home-based visits]

5. Apply technology to reduce the burden of data collection.

Plan (or use an existing one) an application with file sending via email for patients already doing PeakFlow follow-ups if this can replace or complement the certain spirometry (to avoid sending an IDE at home). [Patient #115 with asthma preferring home-based visits]

Discussion

Summary of Findings

In this study, we used a vignette-based survey to solicit patients' preferences for the way RCTs are organized from a large group of patients. We created 6 vignettes based on protocols of real clinical trials. In total, 628 patients with different chronic diseases, ranging from 59 (9.4%) patients with endometriosis to 133 (21.2%) patients with asthma, shared their preferences. Our results highlight the diversity of the preferences and show that if trials are planned according to patients' preferred choices, the likelihood of participating in trials could increase by 30%.

Implications

Our results showed the desire of patients to move from the one-size-fits-all approach of trial participation and tailor the way trials are organized to better suit each patient's condition, such as severity of their diseases, employment status, and distance to the hospital. By allowing flexibility in the way patients participate in a trial, patients who are underrepresented due to barriers, such as employment, income, and distance to hospitals, might be enabled to take part, thus increasing external validity [21,22]. The COVID-19 pandemic has created challenges for clinical trials but also created opportunities to significantly change how trials are conducted. Many clinical trials have adapted to recruit patients online and remotely perform trial visits [23]. For example, the Pragmatic Evaluation of Events And Benefits of Lipid-lowering in Older Adults (PREVENTABLE) trial shipped medications to patients' homes and cognitive assessment was performed at the patient's home by trained research staff [24]. A trial evaluating fluvoxamine to treat COVID-19 was conducted entirely remotely. Patients were recruited online and signed the e-consent form, and data were collected via mail or telephone [25]. These examples show the feasibility of remote trials as an alternative to overcome certain barriers to trial participation. Further research is needed to assess the validity of data collected in a setting outside the research center [26,27]. Nevertheless, other barriers, such as understanding of complex concepts related to clinical trial participation (eg, randomization) or concerns about adverse effects of tested treatments, might not be resolved by remote trials [24]. Thus, discussion to address the concerns of participants and identify the best approach for their participation in trials is necessary.

The patients in our sample also expressed the need to be informed about trial results, preferably during a discussion with a doctor, to get a chance to bring up questions and these being answered directly. This desire is in line with efforts to enhance transparency of trial results for which funders are striving. Further, informing patients about trial results helps them understand the significance of their contribution to science and could encourage them to participate again in future studies [28].

Our study also shows that vignette-based surveys are a useful and innovative design to incorporate perspectives of a large number of patients and public members in research. The vignette is an effective way to communicate the complex concept of clinical trials to patients. In the development of the vignettes for this study, we explained the process and practicalities of taking part in a trial, such as the distance of travel, types and number of examination tests, and the total amount of time for each visit. We cocreated the vignettes with patient representatives to ensure that the structure and language used were comprehensible to the patients. Vignette-based surveys allow patients to express their opinions and ideas without pressure from other stakeholders [29-32]. Vignette-based surveys could also be used to discover patients' perspectives on other aspects of trial design, such as their preferences for comparators, outcomes, and study design.

Limitations

Our study had some limitations. We recruited patients from a patient e-cohort; thus patients in our sample had more experience with the use of the internet and participating in research. In France, 77% of the population has smartphones and 72% of the population has access to the internet [33]. In our study, 427 of 628 (67.9%) patients lived in urban areas and 364 (57.9%) had access to a university hospital in less than 1 hour, which is similar to the French population, with 54.6% of the population

living in urban areas (ie, areas with more than 50,000 inhabitants) [34]. Additionally, the majority of patients (621/628, 98.9%) lived in France; thus their experience with clinical trial participation might be different from that of patients living in other countries in order to be able to generalize the study results. Nevertheless, this study could be adapted to other languages and disseminated to international patient communities. Lastly, due to the hypothetical nature of the vignettes, we cannot exclude the fact that patients might make a different decision in a real-life situation.

Conclusion

We used a vignette-based survey, a new approach, to solicit preferences and ideas to improve RCT organization from a large number of patients. The patients emphasized the need to transform the current one-size-fits-all approach of clinical trial participation.

Acknowledgments

The authors thank Daniel Schanté and Philip Collis for their inputs to the vignettes as patient representatives and their comments on the paper.

This project is part of a Methods in Research on Research (MiRoR) project that received funding from the European Union Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant (agreement no. 676207). The funders had no role in the study design, data collection, and analysis; in the decision to publish; or in the preparation of the manuscript.

Authors' Contributions

Study conception was by VTN, IB, PR, BY, and V-TT; data collection by VTN, V-TT, and IB; data analysis and interpretation by VTN, IB, PR, BY, V-TT, and DS; and writing by VTN, IB, BY, and PR. All authors read and provided feedback on the manuscript. IB is the guarantor. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Transparency declaration: The lead author (IB) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported, that no important aspects of the study have been omitted, and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Conflicts of Interest

VTT is a minority shareholder of SKEZI. PR is a minority shareholder of INATO and SKEZI.

Multimedia Appendix 1
Search strategy for trial protocols.
[DOCX File , 13 KB - jmir_v24i2e29691_app1.docx]

Multimedia Appendix 2
Trial protocols selected for vignette development.
[DOCX File , 25 KB - jmir_v24i2e29691_app2.docx]

Multimedia Appendix 3
Example of a vignette.
[DOCX File , 32 KB - jmir_v24i2e29691_app3.docx]

Multimedia Appendix 4
Nonrespondents' characteristics.
[DOCX File , 13 KB - jmir_v24i2e29691_app4.docx]

References

1. Sully BGO, Julious SA, Nicholl J. A reinvestigation of recruitment to randomised, controlled, multicenter trials: a review of trials funded by two UK funding agencies. Trials 2013 Jun 09;14(1):166 [FREE Full text] [doi: 10.1186/1745-6215-14-166] [Medline: 23758961]
2. Healy P, Galvin S, Williamson PR, Treweek S, Whiting C, Maeso B, et al. Identifying trial recruitment uncertainties using a James Lind Alliance Priority Setting Partnership: the PRioRiTy (Prioritising Recruitment in Randomised Trials) study. Trials 2018 Mar 01;19(1):147 [FREE Full text] [doi: 10.1186/s13063-018-2544-4] [Medline: 29490702]
3. Fletcher B, Gheorghe A, Moore D, Wilson S, Damery S. Improving the recruitment activity of clinicians in randomised controlled trials: a systematic review. BMJ Open 2012 Jan 06;2(1):e000496 [FREE Full text] [doi: 10.1136/bmjopen-2011-000496] [Medline: 22228729]

4. Anderson A, Borfitz D, Getz K. Global public attitudes about clinical research and patient experiences with clinical trials. JAMA Netw Open 2018 Oct 05;1(6):e182969 [FREE Full text] [doi: 10.1001/jamanetworkopen.2018.2969] [Medline: 30646218]
5. Naidoo N, Nguyen VT, Ravaud P, Young B, Amiel P, Schanté D, et al. The research burden of randomized controlled trial participation: a systematic thematic synthesis of qualitative evidence. BMC Med 2020 Jan 20;18(1):6 [FREE Full text] [doi: 10.1186/s12916-019-1476-5] [Medline: 31955710]
6. Dellson P, Nilsson K, Jernström H, Carlsson C. Patients' reasoning regarding the decision to participate in clinical cancer trials: an interview study. Trials 2018 Sep 29;19(1):528 [FREE Full text] [doi: 10.1186/s13063-018-2916-9] [Medline: 30268150]
7. Orri M, Lipset CH, Jacobs BP, Costello AJ, Cummings SR. Web-based trial to evaluate the efficacy and safety of tolterodine ER 4 mg in participants with overactive bladder: REMOTE trial. Contemp Clin Trials 2014 Jul;38(2):190-197 [FREE Full text] [doi: 10.1016/j.cct.2014.04.009] [Medline: 24792229]
8. Jacobs B, Bent S, Tice JA, Blackwell TH, Cummings SR. An internet-based randomized, placebo-controlled trial of kava and valerian for anxiety and insomnia. Medicine (Baltimore) 2005 Jul;84(4):197-207 [FREE Full text] [doi: 10.1097/01.md.0000172299.72364.95] [Medline: 16010204]
9. Galsky MD, Shahin M, Jia R, Shaffer DR, Gimpel-Tetra K, Tsao C, et al. Telemedicine-enabled clinical trial of metformin in patients with prostate cancer. JCO Clin Cancer Inform 2017 Nov(1):1-10. [doi: 10.1200/cci.17.00044]
10. Donnelly S, Reginatto B, Kearns O, Mc Carthy M, Byrom B, Muehlhausen W, et al. The burden of a remote trial in a nursing home setting: qualitative study. J Med Internet Res 2018 Jun 19;20(6):e220 [FREE Full text] [doi: 10.2196/jmir.9638] [Medline: 29921563]
11. Llewellyn-Thomas H, McGreal M, Thiel E, Fine S, Erlichman C. Patients' willingness to enter clinical trials: measuring the association with perceived benefit and preference for decision participation. Soc Sci Med 1991 Jan;32(1):35-42 [FREE Full text] [doi: 10.1016/0277-9536(91)90124-u]
12. Hughes R, Huby M. The application of vignettes in social and nursing research. J Adv Nurs 2002 Feb;37(4):382-386. [doi: 10.1046/j.1365-2648.2002.02100.x] [Medline: 11872108]
13. Bachmann LM, Mühleisen A, Bock A, ter Riet G, Held U, Kessels AG. Vignette studies of medical choice and judgement to study caregivers' medical decision behaviour: systematic review. BMC Med Res Methodol 2008 Jul 30;8(1):50 [FREE Full text] [doi: 10.1186/1471-2288-8-50] [Medline: 18664302]
14. Do-Pham G, Le Cleach L, Giraudeau B, Maruani A, Chosidow O, Ravaud P. Designing randomized-controlled trials to improve head-louse treatment: systematic review using a vignette-based method. J Invest Dermatol 2014 Mar;134(3):628-634 [FREE Full text] [doi: 10.1038/jid.2013.414] [Medline: 24121401]
15. Gould D. Using vignettes to collect data for nursing research studies: how valid are the findings? J Clin Nurs 1996 Jul;5(4):207-212. [doi: 10.1111/j.1365-2702.1996.tb00253.x] [Medline: 8718052]
16. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
17. COMPARE: La communauté de patients pour la recherche. URL: https://compare.aphp.fr/ [accessed 2022-01-12]
18. Tran V, Ravaud P. Collaborative open platform E-cohorts for research acceleration in trials and epidemiology. J Clin Epidemiol 2020 Aug;124:139-148 [FREE Full text] [doi: 10.1016/j.jclinepi.2020.04.021] [Medline: 32380177]
19. Tran V, Riveros C, Péan C, Czarnobroda A, Ravaud P. Patients' perspective on how to improve the care of people with chronic conditions in France: a citizen science study within the ComPaRe e-cohort. BMJ Qual Saf 2019 Nov 23;28(11):875-886 [FREE Full text] [doi: 10.1136/bmjqs-2018-008593] [Medline: 31015376]
20. Tran V, Riveros C, Ravaud P. Patients' views of wearable devices and AI in healthcare: findings from the ComPaRe e-cohort. NPJ Digit Med 2019 Jun 14;2(1):53 [FREE Full text] [doi: 10.1038/s41746-019-0132-y] [Medline: 31304399]
21. Davis TC, Arnold CL, Mills G, Miele L. A qualitative study exploring barriers and facilitators of enrolling underrepresented populations in clinical trials and biobanking. Front Cell Dev Biol 2019;7:74 [FREE Full text] [doi: 10.3389/fcell.2019.00074] [Medline: 31114788]
22. Unger JM, Gralow JR, Albain KS, Ramsey SD, Hershman DL. Patient income level and cancer clinical trial participation: a prospective survey study. JAMA Oncol 2016 Jan 01;2(1):137-139 [FREE Full text] [doi: 10.1001/jamaoncol.2015.3924] [Medline: 26468994]
23. Izmailova ES, Ellis R, Benko C. Remote monitoring in clinical trials during the COVID-19 pandemic. Clin Transl Sci 2020 Sep 08;13(5):838-841 [FREE Full text] [doi: 10.1111/cts.12834] [Medline: 32526077]
24. McDermott MM, Newman AB. Remote research and clinical trial integrity during and after the coronavirus pandemic. JAMA 2021 May 18;325(19):1935-1936. [doi: 10.1001/jama.2021.4609] [Medline: 33885728]
25. Lenze EJ, Mattar C, Zorumski CF, Stevens A, Schweiger J, Nicol GE, et al. Fluvoxamine vs placebo and clinical deterioration in outpatients with symptomatic COVID-19: a randomized clinical trial. JAMA 2020 Dec 08;324(22):2292-2300 [FREE Full text] [doi: 10.1001/jama.2020.22760] [Medline: 33180097]
26. Tunji-Ajayi P, Duncan EM, Gillies K. An embedded mixed-methods study highlighted a lack of discussions on retention in clinical trial consultations. J Clin Epidemiol 2020 Jul;123:49-58 [FREE Full text] [doi: 10.1016/j.jclinepi.2020.03.011] [Medline: 32229250]

27. von Niederhäusern B, Saccilotto R, Schädelin S, Ziesenitz V, Benkert P, Decker M, et al. Validity of mobile electronic data capture in clinical studies: a pilot study in a pediatric population. BMC Med Res Methodol 2017 Dec 08;17(1):163 [FREE Full text] [doi: 10.1186/s12874-017-0438-x] [Medline: 29216831]
28. National Health Service Health Research Authority. Make It Public: Transparency and Openness in Health and Social Care Research. URL: https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/research-transparency/make-it-public-transparency-and-openness-health-and-social-care-research/ [accessed 2022-01-12]
29. Boivin A, Richards T, Forsythe L, Grégoire A, L'Espérance A, Abelson J, et al. Evaluating patient and public involvement in research. BMJ 2018 Dec 06;363:k5147. [doi: 10.1136/bmj.k5147] [Medline: 30522999]
30. Wicks P. Could digital patient communities be the launch pad for patient-centric trial design? Trials 2014 May 15;15(1):172 [FREE Full text] [doi: 10.1186/1745-6215-15-172] [Medline: 24885434]
31. Karazivan P, Dumez V, Flora L, Pomey M, Del Grande C, Ghadiri DP, et al. The patient-as-partner approach in health care. Acad Med 2015;90(4):437-441 [FREE Full text] [doi: 10.1097/acm.0000000000000603]
32. Mader LB, Harris T, Kläger S, Wilkinson IB, Hiemstra TF. Inverting the patient involvement paradigm: defining patient led research. Res Involv Engagem 2018 Jul;4:21 [FREE Full text] [doi: 10.1186/s40900-018-0104-4] [Medline: 30002875]
33. Baromètre du numérique: Enquête sur la diffusion des technologies de l'informationet de la communication dans la société française en 2019. 2019. URL: https://www.arcep.fr/uploads/tx_gspublication/rapport-barometre-num-2019.pdf [accessed 2022-01-12]
34. Costemalle V. Toujours plus d'habitants dans les unités urbaines. URL: https://www.insee.fr/fr/statistiques/4806684 [accessed 2022-01-12]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
ComPaRe: La communauté de patients pour la recherche
PREVENTABLE: Pragmatic Evaluation of Events And Benefits of Lipid-lowering in Older Adults
RCT: randomized controlled trial

Edited by R Kukafka; submitted 16.04.21; peer-reviewed by R Krukowski, T Davis, R Mpofu; comments to author 24.05.21; revised version received 04.06.21; accepted 21.11.21; published 01.02.22.

Please cite as:
Nguyen VT, Ravaud P, Tran VT, Young B, Boutron I
Patients' Perspectives on Transforming Clinical Trial Participation: Large Online Vignette-based Survey
J Med Internet Res 2022;24(2):e29691
URL: https://www.jmir.org/2022/2/e29691
doi:10.2196/29691
PMID:35103603

Original Paper

Identification of Social Engagement Indicators Associated With Autism Spectrum Disorder Using a Game-Based Mobile App: Comparative Study of Gaze Fixation and Visual Scanning Methods

Maya Varma[1], BS; Peter Washington[2], MS; Brianna Chrisman[2], BS; Aaron Kline[3], BS; Emilie Leblanc[3], MS; Kelley Paskov[4], MS; Nate Stockham[5], MS; Jae-Yoon Jung[3], PhD; Min Woo Sun[4], BS; Dennis P Wall[3], PhD

[1]Department of Computer Science, Stanford University, Stanford, CA, United States

[2]Department of Bioengineering, Stanford University, Stanford, CA, United States

[3]Department of Pediatrics and Biomedical Data Science, Stanford University, Stanford, CA, United States

[4]Department of Biomedical Data Science, Stanford University, Stanford, CA, United States

[5]Department of Neuroscience, Stanford University, Stanford, CA, United States

Corresponding Author:
Dennis P Wall, PhD
Department of Pediatrics and Biomedical Data Science
Stanford University
1265 Welch Road
Stanford, CA, 94304
United States
Phone: 1 650 497 9214
Email: dpwall@stanford.edu

Abstract

Background: Autism spectrum disorder (ASD) is a widespread neurodevelopmental condition with a range of potential causes and symptoms. Standard diagnostic mechanisms for ASD, which involve lengthy parent questionnaires and clinical observation, often result in long waiting times for results. Recent advances in computer vision and mobile technology hold potential for speeding up the diagnostic process by enabling computational analysis of behavioral and social impairments from home videos. Such techniques can improve objectivity and contribute quantitatively to the diagnostic process.

Objective: In this work, we evaluate whether home videos collected from a game-based mobile app can be used to provide diagnostic insights into ASD. To the best of our knowledge, this is the first study attempting to identify potential social indicators of ASD from mobile phone videos without the use of eye-tracking hardware, manual annotations, and structured scenarios or clinical environments.

Methods: Here, we used a mobile health app to collect over 11 hours of video footage depicting 95 children engaged in gameplay in a natural home environment. We used automated data set annotations to analyze two social indicators that have previously been shown to differ between children with ASD and their neurotypical (NT) peers: (1) gaze fixation patterns, which represent regions of an individual's visual focus and (2) visual scanning methods, which refer to the ways in which individuals scan their surrounding environment. We compared the gaze fixation and visual scanning methods used by children during a 90-second gameplay video to identify statistically significant differences between the 2 cohorts; we then trained a long short-term memory (LSTM) neural network to determine if gaze indicators could be predictive of ASD.

Results: Our results show that gaze fixation patterns differ between the 2 cohorts; specifically, we could identify 1 statistically significant region of fixation (P<.001). In addition, we also demonstrate that there are unique visual scanning patterns that exist for individuals with ASD when compared to NT children (P<.001). A deep learning model trained on coarse gaze fixation annotations demonstrates mild predictive power in identifying ASD.

Conclusions: Ultimately, our study demonstrates that heterogeneous video data sets collected from mobile devices hold potential for quantifying visual patterns and providing insights into ASD. We show the importance of automated labeling techniques in generating large-scale data sets while simultaneously preserving the privacy of participants, and we demonstrate that specific social engagement indicators associated with ASD can be identified and characterized using such data.

(J Med Internet Res 2022;24(2):e31830) doi:10.2196/31830

KEYWORDS

mobile health; autism spectrum disorder; social phenotyping; computer vision; gaze; mobile diagnostics; pattern recognition; autism; diagnostic; pattern; engagement; gaming; app; insight; vision; video

Introduction

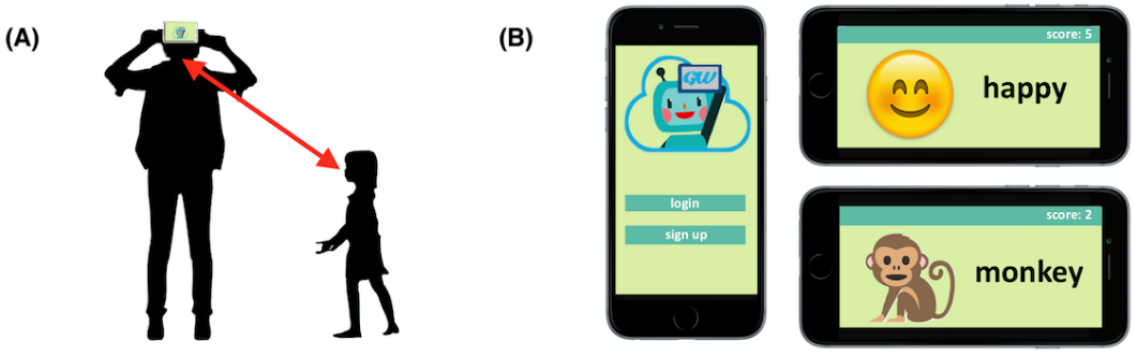
Background

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by social impairments, communication difficulties, and restricted and repetitive patterns of behavior. Currently, 1 in 44 children in the United States have been diagnosed with ASD, with males 4 times more likely to be affected than females [1,2]. ASD usually manifests in infants and children and presents a wide range of symptoms that vary in intensity from person to person. The heterogeneity of ASD presents a major diagnostic challenge, with clinicians typically employing a combination of lengthy parent questionnaires and clinical observation to evaluate children.

Standard diagnostic mechanisms for ASD are often accompanied by a range of issues that result in long waiting times for results [3-5]. However, in recent years, significant strides have been made in the fields of computer vision and mobile technology, giving rise to the possibility of using home videos of a child's natural behaviors to identify characteristics linked with ASD and enable a more accurate and timely diagnosis [6].

We previously created a mobile app called GuessWhat, which yields video data of children engaged in socially motivated gameplay with parents in a natural home environment [7-13]. The app presents a charades game, encouraging kids to act out a series of given prompts, such as emotions, sports, or chores. During a game, parents will open the GuessWhat app and place the smartphone on their foreheads, with the front-facing camera pointing at the child; the child then proceeds to act out the prompt displayed on the device while the parent attempts to predict the answer, as shown in Figure 1. The game ends when the 90-second time limit is exceeded. At this point, the parent can view the video recording of the child and is then given the option to share this data with our research team.

Figure 1. GuessWhat Mobile app. (A) The parent places the mobile phone in a fixed location, allowing the recording of a semistructured gameplay video. (B) The children are presented with a variety of charades prompts, such as emotions and animals.



The data collection pipeline employed by GuessWhat provides several benefits that make the obtained information amenable to computational analysis. First, although children are performing varied tasks in diverse environments, GuessWhat videos encourage inherent structure, with factors such as the position of the phone camera, location of the child relative to the camera, and game-based social interactions between the parent and child remaining generally consistent between videos. In addition, as children are in a home environment and are unencumbered by bulky hardware such as eye trackers or head mounts, they can interact with their parents and surroundings in a natural manner. As a result, we hypothesize that computer vision algorithms can be designed to monitor socially motivated facial engagement in children during gameplay, allowing effective identification of behaviors, eye contact events, and social interactions potentially correlated with the ASD phenotype.

In this work, we used computational techniques to analyze these videos and identify differences in social interactions between children with ASD and neurotypical (NT) children. We specifically analyzed 2 common social engagement signals that are included in standard clinical diagnostic instruments and can be identified through computer vision methodologies: (1) gaze fixation patterns, which represent the regions of an individual's visual focus and (2) visual scanning methods, which refer to the ways in which individuals scan their surrounding environment. We performed these tasks without sharing participant videos or private patient information with human annotators.

Ultimately, the development of this system can help improve diagnosis of ASD through automated detection of impaired social interactions, mitigating the problems associated with limited diagnostic resources for neurodevelopmental disorders, especially in regions where access to care is limited [14]. This work also demonstrates the usefulness of game-based approaches and automated labeling methods in preserving privacy, generating large diagnostic data sets, and improving human understanding of complex conditions.

Prior Work

Researchers have demonstrated the usefulness of video data in providing diagnostic insights into gaze and engagement behaviors associated with ASD. Prior work can generally be divided into three categories: (1) manual annotation methods, (2) eye-tracking systems, and (3) use of structured environments.

Manual Annotation Methods

Some studies have used human annotators to label social interaction and engagement information in video frames. Several prior works, such as those by Tariq et al and Leblanc et al, performed manual annotation of behavioral features in home videos, which enabled the creation of classifiers that could identify ASD with high accuracy [15-20]. Chorianopoulou et al collected structured home videos from participants and had expert annotators label the data set with the actions, emotions, gaze fixations, utterances, and overall level of engagement in each video; this information was then used to train a classifier to identify specific engagement features that could be correlated with ASD [21]. Rudovic et al trained a large and generalizable neural network to estimate engagement in children with ASD from different cultural backgrounds [22]. Engagement labels were manually annotated by trained individuals. Although these methods enable the creation of human-vetted, accurate data sets, such approaches require large numbers of trained annotators when implemented on a large scale, which is expensive and time-consuming. In addition, these techniques may compromise the privacy of participants by providing annotators with access to video footage, although some methods have been developed to address privacy concerns with crowdsourced annotations [23,24].

Eye-Tracking Systems

Several studies have used eye trackers to identify patterns in gaze and engagement behaviors that may be indicative of ASD or other developmental conditions [25-28]. Pusiol et al showed that deep learning models trained on data collected from a head-mounted eye tracker and camera could be used to classify idiopathic developmental disorder and fragile X syndrome with high precision [29]. Similarly, Riby et al used eye trackers to show that individuals with ASD had atypical gaze patterns when watching movies and cartoons [30]. To counteract artificial movements often associated with facial eye trackers, Noris et al developed a nonintrusive eye-tracking device mounted on a hat that recorded a child's interactions with an interviewer; the study concluded that children with ASD were more inclined to look downward during social interaction than their NT peers [31]. Despite the accuracy and quality of gaze data collected from such systems, eye trackers require custom hardware that can often be expensive and inaccessible, especially for individuals living in resource-limited regions. As a result, these approaches are unlikely to be accessible to the general population.

Use of Structured Environments

Hashemi et al explored the use of computer vision algorithms to identify behaviors associated with ASD [32]. A trained clinician administered a series of predefined, structured tasks involving toys and other visual stimuli, while a video camera captured footage of the child's response. A computer vision system that analyzed the child's body orientation and facial movement was able to evaluate the child's engagement with high accuracy. Similarly, Chang et al used the front-facing camera of a mobile device to capture gaze scanning patterns as children watched strategically designed short movies [33]. Automated computer vision techniques were then used to identify differences in gaze patterns between children with ASD and NT individuals. Egger et al also used mobile phones to collect videos of ASD and NT children engaging with short movies. Visual stimuli in movies were carefully designed based on neuroscience principles, and children's emotional and behavioral responses were computationally analyzed [34]. These works demonstrate effective methods for analyzing engagement patterns without the use of manual annotations or external eye-tracking hardware; however, these studies were conducted with highly structured tasks (eg, carefully selected movies and toys) and controlled environmental factors (eg, Hashemi et al and Chang et al controlled the room lighting and distance of the camera from the participant's face). As a result, the ability of these techniques to translate to natural nonclinical environments and unstructured tasks remains to be explored. In addition, these works do not evaluate engagement and behaviors in social situations.

Our Contributions

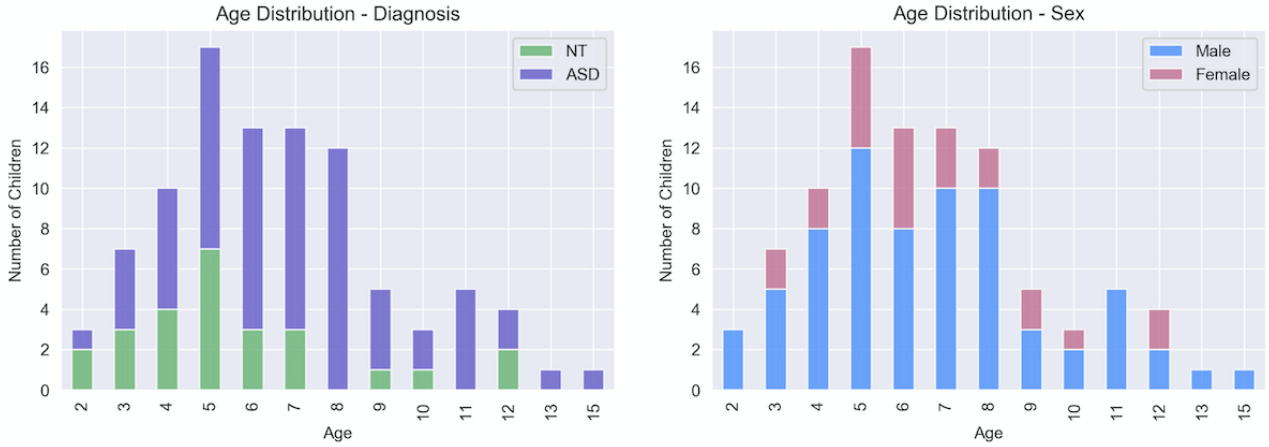
To the best of our knowledge, this is the first study that attempts to obtain diagnostic insights into ASD from social gameplay videos without the use of eye-tracking hardware, manual frame-level annotations, and structured scenarios or environments. We show that semistructured gameplay videos collected on mobile devices reveal specific regions of gaze fixation as well as visual scanning patterns that differ between individuals with ASD and NT children during social gameplay. With further research and development, our system can be deployed as a diagnostic tool in diverse settings on a large scale.

Methods

Data Collection

We used the GuessWhat mobile app to collect videos of children engaged in gameplay with a parent. Participants were recruited using social media advertisements and research email lists maintained by the study team. Approximately 1000 individuals proceeded to download the GuessWhat app, and we collected 449 videos from 95 children for this study. The participants ranged in age from 2 to 15 years and included 68 children (15 females, 53 males) diagnosed with ASD as well as 27 NT children (9 females, 18 males). Each child contributed a mean of 4.7 videos (SD 7.3), resulting in a total data set size of 1,084,267 individual frames and 11.1 hours of footage, presented in Figure 2. All parents consented to share their videos with our research team and completed a survey to provide the age, sex, and diagnostic status of their children.

Figure 2. Data set information. These graphs show the breakdown of our data set by age, diagnosis, and sex. In our data set, 1 NT male failed to provide his age, and this information has been excluded from this figure. ASD: autism spectrum disorder; NT: neurotypical.



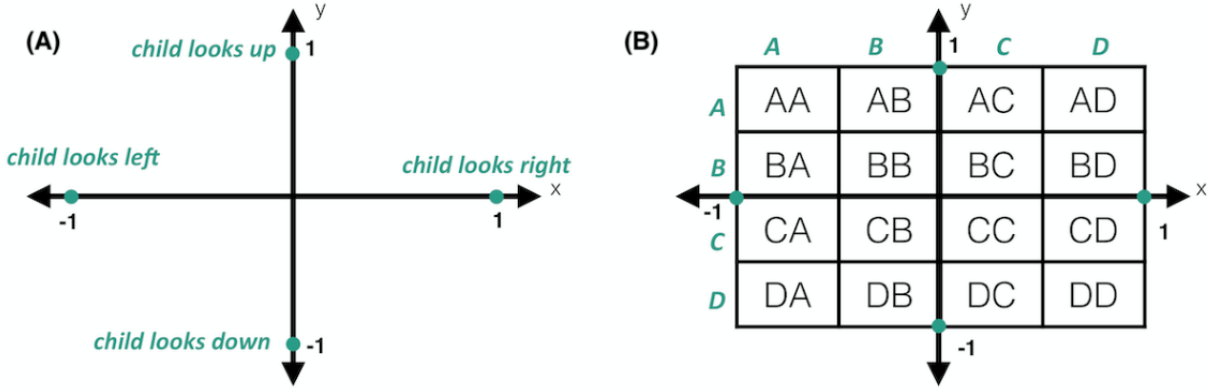
Data Preprocessing

Although the semistructured format of our video data set presents numerous advantages, home videos are naturally heterogeneous in quality; this results in several challenges that must be addressed prior to computational analysis. Specifically, excessive camera movement and poor lighting conditions rendered some frames in our data set too blurry for use. Moreover, other adults or siblings would often join in gameplay, resulting in multiple faces in the frame and making identification of the participating child challenging. Another major challenge arises from the lack of fine-grained annotations and ground truth labels; although the lack of eye-tracking hardware enables natural child motions and interactions, this also results in a lack of calibration information for obtaining accurate gaze locations.

We began our analysis with extensive quality control and data preprocessing. To preserve privacy, we annotated our data set solely using computational methods. We first used Amazon Rekognition, a powerful off-the-shelf computer vision platform developed by Amazon, to perform noisy labeling of key features in each still frame, including 30 facial landmarks and facial bounding boxes. Frames with 0 or greater than 2 faces were removed from the data set. We then used an open-source facial landmark annotation platform called OpenFace to obtain automated estimates of gaze directions [35,36]. Each frame with an identifiable face was assigned a coordinate pair (x,y) representing the direction of the individual's gaze. The value of x ranges from –1 (indicating a leftward gaze) to 1 (indicating a rightward gaze); similarly, the value of y ranges from –1 (indicating a downward gaze) to 1 (indicating an upward gaze), as shown in Figure 3. As these coordinates were assigned with respect to the smartphone camera, a frame in which an individual is gazing straight ahead into the camera is assigned a coordinate pair of (0,0). If the OpenFace model demonstrated low confidence in gaze estimation values (defined as confidence below 75%) because of occluded eyes or insufficient image quality, the frame was removed from the data set; as a result, we expected the final annotations to be of high quality, but the presence of some noise and incorrect labels was to be expected. This procedure resulted in a total of 619,620 annotated frames, representing 520,536 frames from children with ASD and 99,084 frames from NT children.

Finally, to discretize gaze annotation data, we divided the coordinate map into 16 distinct areas of interest (AOIs), as shown in Figure 3. All gaze coordinates that fell within the bounds of a particular AOI were grouped together. Such an approach allowed us to identify trends in an individual's gaze fixations and scanning patterns.

Figure 3. Gaze annotations. (A) Gaze coordinates range between –1 and 1 on the x- and y-axes. (B) To categorize gaze coordinates into discrete regions, we divided the gaze map into 16 buckets. Each area of interest is labeled with corresponding row and column letters.



Differential Pattern Analysis

Gaze Fixation Patterns

Gaze fixation, which occurs when one's gaze is held on a single target for an extended period, plays an important role in social interaction by signaling communicative intent and enabling interpersonal relationships. In a dyadic social interaction, individuals usually fixate their gaze on the target's eyes. However, individuals with ASD often face difficulty with maintaining eye contact and instead tend to focus their visual attention on other regions of the target's face. Several studies involving eye trackers and visual stimuli have shown that children with ASD tend to fixate on the mouth or other body parts; this has even been observed in children aged as young as 2 to 6 months who were later diagnosed with ASD [37-39]. Eye contact avoidance, which is explicitly examined in standard clinical diagnostic examinations, can result in decreased facial identification and social engagement.

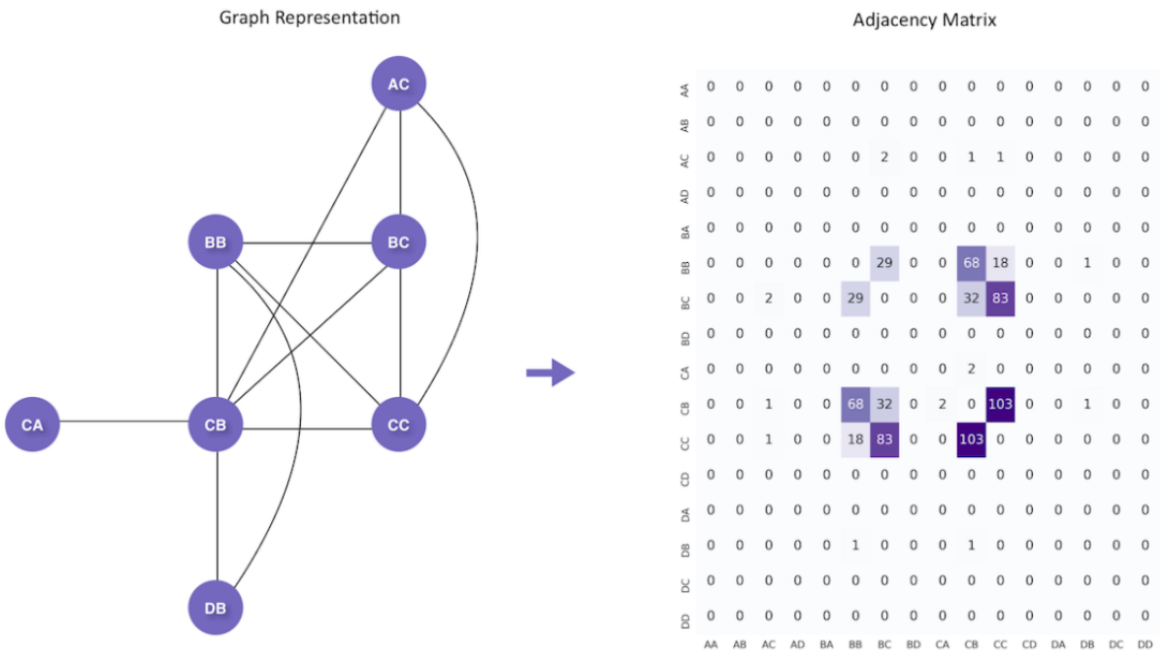
To determine the gaze fixation patterns of individuals during a single 90-second game, we used the coarse gaze annotations obtained from our preprocessed data set. For each video in our data set, we computed the percentage of time that the child fixated his or her gaze on each of the 16 predefined AOIs. A 2-sided permutation test was used at every AOI to identify statistically significant differences between the ASD and NT populations, with the null hypothesis that the fixation times for both populations followed an equivalent distribution; we calculated the difference in the mean fixation times for 100,000 rearrangements of the 2 groups. Bonferroni correction was applied to account for multiple hypothesis tests. It is important to note that because the AOIs are correlated, the Bonferroni correction is extremely stringent and will reduce the likelihood of Type 1 errors.

Visual Scanning Patterns

Humans tend to transition their gaze between various objects in their environments when encountering visual stimuli, a phenomenon called visual scanning. The patterns and frequencies with which humans scan their surroundings can provide insight into how individuals process the world around them. In the context of social interaction, prior research has shown that individuals with ASD vary in the way that they scan a target's facial landmarks during a social scenario, which may contribute to difficulty with interpreting emotional or nonverbal cues. This was shown by Pelphrey et al, who demonstrated that when presented with images of faces, NT individuals typically transitioned their gaze between core features, such as the eyes and nose, whereas individuals with ASD appeared to scan nonfeature areas of the face, such as the forehead and cheeks [40]. A similar study conducted by Chawarska and Shik on toddlers corroborated these findings, providing evidence of atypical scanning patterns in children with ASD when compared to their age-matched NT peers [41]. Understanding these patterns can reveal differences in the way that individuals with ASD process visual stimuli and interact in social situations.

Modeling gaze transition patterns as a graph problem can provide insight into the regions that children focus on while scanning their environments [42]. For each 90-second video of gameplay, we constructed a network consisting of 16 nodes n_{AA}, n_{AB}, …, n_{DC}, n_{DD}, with each node representing a predefined AOI. When a child shifts his or her gaze between locations on the 16-AOI gaze map, an undirected edge $e=(n_i, n_j)$ is drawn between the 2 corresponding nodes. Edges are weighted by the number of transitions that occur during the game. The graph can then be converted to a 16 × 16 adjacency matrix, as depicted in Figure 4.

Figure 4. Graph model of gaze transitions. We modeled the gaze transitions in each gameplay video as a graph, which was then used to generate a 16 × 16 adjacency matrix.



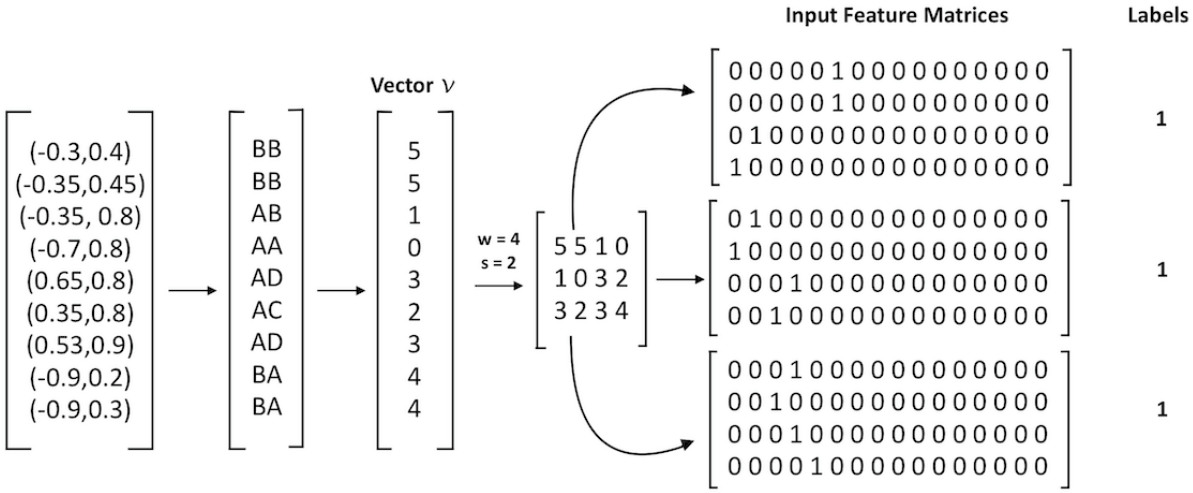
We computed adjacency matrices for all gameplay videos and normalized each matrix by dividing each entry by the total number of transitions. Then, we computed the average of all matrices associated with the NT individuals in our data set, resulting in a single 16 × 16 matrix depicting the mean percentage of transitions occurring between each pair of AOIs in a single game. This process was repeated for the gameplay videos associated with the ASD cohort. We conducted 2-sided permutation tests at each location in the transition matrix to determine if there were significant differences in the transition types between the 2 groups.

Deep Learning Model

Next, we used deep learning techniques to measure the predictive power of gaze fixation patterns. We began by converting fixation data points into feature matrices that could serve as the input to our classifiers. We first extracted the sequence of gaze coordinates from each video using the coarse annotation procedure described in the previous section. This resulted in a vector of n ordered pairs (x,y) for every video, where n represents the number of valid frames in the video, and x and y are the gaze fixation coordinates ranging from –1 to 1. We then matched each ordered pair with its associated AOI, as demonstrated in Figure 3. This yielded a vector of n AOIs, representing the regions of the gaze map that each individual fixated on during a game. Next, each of the 16 predefined AOIs was assigned a number from 0 to 15 in alphabetical order, with 0 representing AA and 15 representing DD; this formed a vector of n integers, which we will refer to as v.

We used a sliding window approach to divide v into separate vectors using 2 predefined parameters, namely window and shift. The window parameter w represents the number of frames included in a single feature vector; in our experiments, this value ranged from 50 to 500 frames, which roughly corresponds to 2 to 20 seconds of video content. The shift parameter s defines the number of elements by which the window slides between feature vectors, and we experimented with shift values between 10 and 100. These parameters allowed us to extract feature vectors from v consisting of w elements, with vectors separated by exactly s frames; note that if $s<w$, vectors will contain overlapping elements. Finally, we converted each w vector into a $w \times 16$ feature matrix, with each AOI integer encoded by a one-hot vector. A demonstrative example is shown in Figure 5.

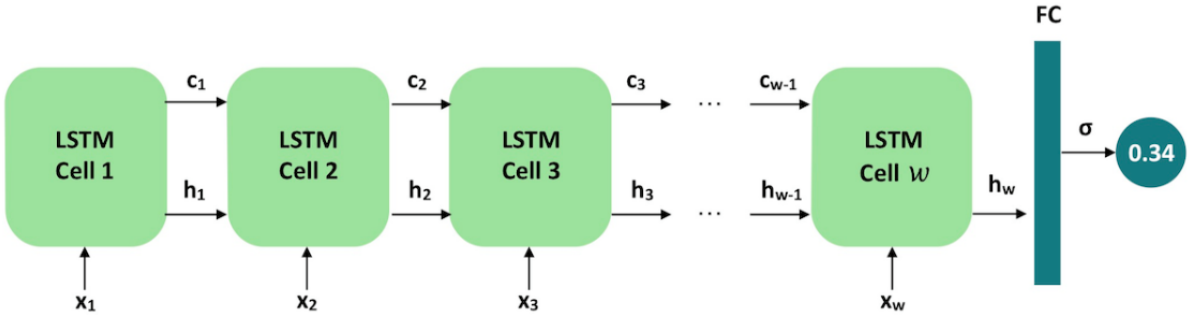
Figure 5. Gaze fixation feature representation. In this demonstrative example, we begin with a video consisting of 9 frames. Gaze coordinates are matched with corresponding area of interest (AOI) regions. Using a window of 4 and a shift value of 2 divides vector v into 3 feature vectors. Each feature vector is then one-hot encoded. All input feature matrices are assigned the same label.



We then used deep learning models to determine if gaze fixation patterns could be predictive of ASD. We assigned 324 videos (275 ASD, 49 NT) in our data set to the training set, 71 videos (62 ASD, 9 NT) to the validation set, and 54 videos (43 ASD, 11 NT) to the held-out test set, ensuring that all videos corresponding to a single child were assigned to the same set. Input feature matrices were constructed using the approach described above. A binary label $l \in \{0,1\}$ was assigned to each matrix to represent the diagnosis of the child in the associated video, with 1 representing the presence of ASD.

To exploit the temporal nature of our data set, we used long short-term memory (LSTM) networks, which are a type of recurrent neural network that can model long-term dependencies. A $w \times 16$ feature matrix served as the input to an LSTM model with w cells; each cell accepted a one-hot encoded 16-feature vector as the input. We used the Adam optimizer with a learning rate of 0.001, a batch size of 5, and a weighted binary cross-entropy loss function. The last cell of the LSTM network was connected to a fully connected layer with a single class output followed by a sigmoid nonlinearity; this resulted in a final value ranging between 0 and 1. This value was rounded to the closest integer to determine the final prediction, as observed in Figure 6.

Figure 6. Model architecture. The model consists of a long short-term memory network with w cells. Each cell accepts a one-hot vector of size 16, represented in the figure by x_i, and outputs a cell state c_i and a hidden state h_i. The final cell is connected to a fully connected layer, which generates a single class output. FC: fully connected layer; LSTM: long short-term memory.



Finally, to characterize model performance, we report four metrics: macroaveraged recall, macroaveraged precision, weighted-average recall, and weighted-average precision. As our data set exhibits class imbalance with cases outnumbering controls, these metrics provide the most accurate representation of model performance. Macroaveraged statistics compute the arithmetic mean of performance on each class, whereas weighted-average statistics compute the weighted mean. We performed all parameter experiments on our validation set and evaluated our final best-performing models on the held-out test set.

Ethics Approval

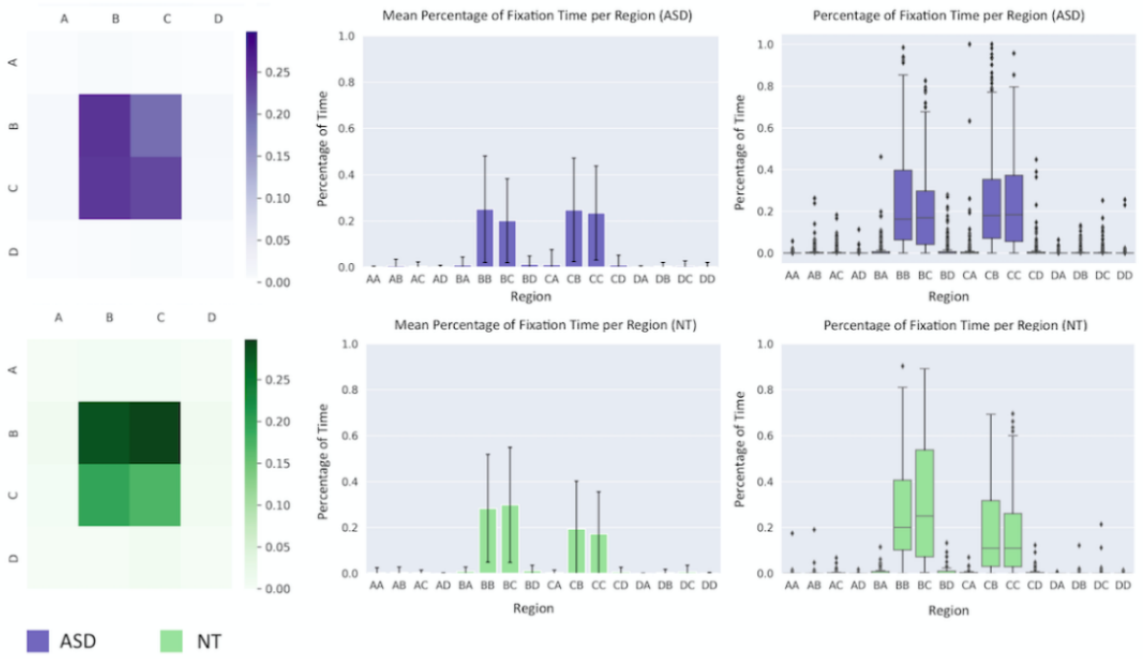
This study was approved by the Stanford Institutional Review Board (eProtocol number: 39562).

Results

Gaze Fixation Patterns Differ Between ASD and NT

We first analyzed gaze fixation patterns to determine if regions of focus differ between children with ASD and NT children during a single 90-second game. Coarse gaze annotations, which were obtained using the automated labeling procedure described in the Methods section, were grouped into 16 AOIs, and the percentage of time that the child fixated on each region was computed. Figure 7 shows the mean percentage of time that the ASD and NT cohorts fixated on each AOI during a game. As shown by the heat maps, children mostly fixated on the 4 central locations BB, BC, CB, and CC, which are located closest to the camera of the mobile phone. The distributions show that several differences exist between the 2 populations; children with ASD were most likely to fixate on locations BB and CB, whereas NT children spent much of the 90-second game focusing on locations BB and BC. We conducted a 2-sided permutation test at each AOI with 100,000 permutations of the data, setting a Bonferroni-corrected significance threshold of 0.0031 to account for the 16 hypothesis tests. A significant difference in fixation distributions between the 2 cohorts was observed at location BC (P<.001).

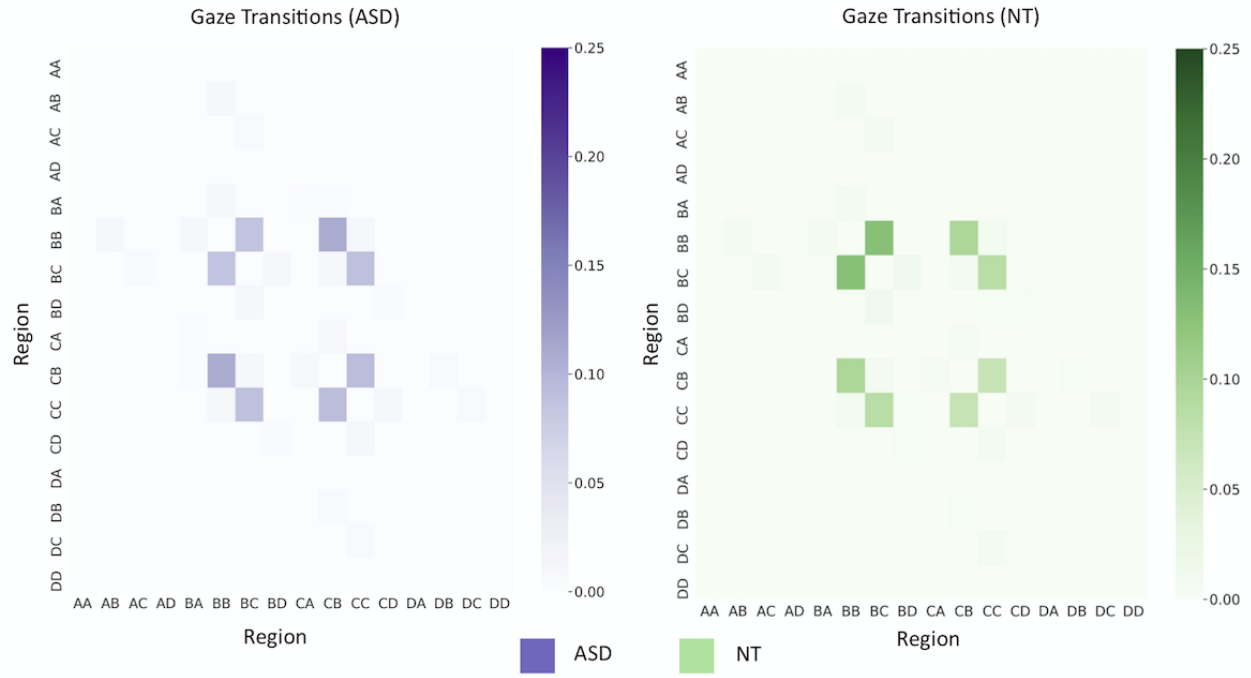
Figure 7. Gaze fixation results. The heat maps located at the upper left and lower left show the mean percentage of time that an individual fixated his or her gaze on each area of interest (AOI). The bar charts and the box and whisker plots show the distribution of fixation times across all videos. ASD: autism spectrum disorder; NT: neurotypical.



Visual Scanning Patterns Differ Between ASD and NT

Next, we used graph methods to analyze the ways in which participants scanned their environments during gameplay. We modeled the gaze transitions in each gameplay video as a network and computed the mean adjacency matrices for the ASD and NT populations, which are shown in Figure 8; a cell of the matrix in row *i* and column *j* represents the mean percentage of gaze transitions in a single 90-second game that occur between AOI *i* and AOI *j*. We conducted permutation tests with 100,000 permutations at each of the 61 nonzero, unique locations in the adjacency matrices; as the matrix is symmetric, the distributions for each distinct transition pair were tested for significance exactly once. We then used a Bonferroni-corrected significance threshold of 0.0008 to account for 61 hypothesis tests. Our results show that a significant difference exists in the percentage of gaze transitions between regions BB and BC (P<.001). As shown by the heat maps in Figure 8, 9.4% of the gaze transitions made by an individual with ASD occur between BB and BC; however, for NT children, 13% of the gaze transitions made during a 90-second game occur between BB and BC.

Figure 8. Gaze transition heat maps. These heat maps show the percentage of gaze transitions that occur between each pair of AOIs during a 90-second game. AOI: area of interest; ASD: autism spectrum disorder; NT: neurotypical.



Gaze Fixation Patterns Provide Mild Predictive Power

We measured the classification performance of models trained on gaze fixation patterns. Gaze fixation coordinates were encoded as one-hot vectors and passed as input to an LSTM network, which generated a single class output representing the likelihood of ASD. LSTM models were trained with a range of window and shift parameter values and evaluated on the validation set. Our results from the validation set allowed us to identify our top 3 models, which were trained with parameters (1) w=100, s=10; (2) w=200, s=10; and (3) w=500, s=10. These networks were then evaluated on the held-out test set. In Table 1, we provide precision and recall values for an LSTM model trained with those values of the window w and shift s that achieved the best performance on the validation set.

Table 1. Classifier performance on held-out test set with gaze fixation features.

Window (w)	Shift (s)	Macroaveraged recall	Macroaveraged precision	Weighted-average recall	Weighted-average precision
100	10	0.598	0.595	0.656	0.661
200	10	0.561	0.577	0.662	0.635
500	10	0.576	0.577	0.625	0.624

The model with parameters w=100 and s=10 demonstrated the best performance. Macroaveraged statistics are lower than weighted-average statistics, suggesting that the accuracy of prediction differs between the 2 classes. In summary, the results suggest that gaze fixation patterns can provide mild predictive power.

Discussion

In this study, we used computational techniques to analyze home videos and obtain diagnostic insights into ASD. We collected a large data set of semistructured videos featuring children engaged in gameplay with a parent, and we analyzed 2 key markers of social engagement that have been shown to differ

between children with ASD and their NT peers: (1) gaze fixation and (2) visual scanning. For each marker, we identified statistically significant differences between the 2 cohorts and demonstrated that this information could be useful in identifying the presence of ASD.

Our study demonstrates the potential that mobile tools hold for quantifying visual patterns and providing insights into ASD. Despite the presence of high heterogeneity and varying quality in our data set, the automated labeling techniques and deep learning classifiers used in this work were able to extract usable signals and identify differences in gaze fixation and visual scanning patterns between the 2 cohorts. These methods also enabled us to preserve participant privacy by avoiding the use of human annotators. Our findings support prior works that have identified social and visual engagement differences between individuals with ASD and NT individuals [37-41], and we demonstrate here that these variations can be identified using mobile tools. In contrast to previous video-based diagnostic approaches, we demonstrate that diagnostic insights can be obtained without the use of manual annotation methods, eye-tracking systems, or structured environments.

This work has some limitations. First, due to the class imbalance in our data set, the predictive accuracy of ASD differs from that of the control individuals; this is reflected in Table 1, which shows variations between macroaveraged statistics and weighted-average statistics. Additional data set augmentations will be necessary to correct this issue in future. In addition, due to camera motion and variation in the location of the smartphone relative to the parent's face, the gaze fixation maps are difficult to interpret qualitatively, and AOIs cannot be definitely matched to a parent's specific facial regions.

Future directions for this work include expanding the size of the experimental population; analyzing additional motion-based features in gameplay videos, such as limb movements and coordination; performing qualitative human-centered investigations or pragmatic randomized controlled trials to evaluate clinical usability; and evaluating the real-world diagnostic capabilities of our approach across diverse environmental settings [43-47].

Overall, this study demonstrates the usefulness of game-based mobile apps and heterogeneous video data sets in aiding in the diagnosis of ASD. With further research and development, the system described in this work can ultimately serve as a low-cost and accessible diagnostic tool for a global population.

Acknowledgments

This work was supported in part by funds to DPW from the National Institutes of Health (1R01EB025025-01, 1R01LM013364-01, 1R21HD091500-01, 1R01LM013083), the National Science Foundation (Award 2014232), The Hartwell Foundation, Bill and Melinda Gates Foundation, Coulter Foundation, Lucile Packard Foundation, Auxiliaries Endowment, the ISDB Transform Fund, the Weston Havens Foundation, and program grants from Stanford’s Human Centered Artificial Intelligence Program, Precision Health and Integrated Diagnostics Center, Beckman Center, Bio-X Center, Predictives and Diagnostics Accelerator, Spectrum, Spark Program in Translational Research, MediaX, and from the Wu Tsai Neurosciences Institute's Neuroscience: Translate Program. We also acknowledge generous support from David Orr, Imma Calvo, Bobby Dekesyer and Peter Sullivan. MV is supported by the Knight-Hennessy Scholars program at Stanford University and the National Defense Science and Engineering Graduate (NDSEG) Fellowship. PW would like to acknowledge support from Mr. Schroeder and the Stanford Interdisciplinary Graduate Fellowship (SIGF) as the Schroeder Family Goldman Sachs Graduate Fellow.

Conflicts of Interest

DPW is the founder of Cognoa.com. This company is developing digital health solutions for pediatric care. AK works as a part-time consultant with Cognoa.com. All other authors declare no conflict of interests.

References

1. Fombonne E. Editorial: the rising prevalence of autism. J Child Psychol Psychiatry 2018 Jul;59(7):717-720. [doi: 10.1111/jcpp.12941] [Medline: 29924395]
2. Maenner MJ, Shaw KA, Baio J, Washington A, Patrick M, DiRienzo M, et al. Prevalence of autism spectrum disorder among children aged 8 years-autism and developmental disabilities monitoring network, 11 sites, United States, 2016. MMWR Surveill Summ 2020 Mar;69(4):1-12 [FREE Full text] [doi: 10.15585/mmwr.ss6904a1] [Medline: 32214087]
3. Bisgaier J, Levinson D, Cutts DB, Rhodes KV. Access to autism evaluation appointments with developmental-behavioral and neurodevelopmental subspecialists. Arch Pediatr Adolesc Med 2011 Jul;165(7):673-674. [doi: 10.1001/archpediatrics.2011.90] [Medline: 21727283]
4. Gordon-Lipkin E, Foster J, Peacock G. Whittling down the wait time: exploring models to minimize the delay from initial concern to diagnosis and treatment of autism spectrum disorder. Pediatr Clin North Am 2016 Oct;63(5):851-859 [FREE Full text] [doi: 10.1016/j.pcl.2016.06.007] [Medline: 27565363]
5. Crane L, Chester JW, Goddard L, Henry LA, Hill E. Experiences of autism diagnosis: a survey of over 1000 parents in the United Kingdom. Autism 2016 Feb;20(2):153-162. [doi: 10.1177/1362361315573636] [Medline: 25810370]

6. Washington P, Park N, Srivastava P, Voss C, Kline A, Varma M, et al. Data-driven diagnostics and the potential of mobile artificial intelligence for digital therapeutic phenotyping in computational psychiatry. Biol Psychiatry Cogn Neurosci Neuroimaging 2020 Aug;5(8):759-769 [FREE Full text] [doi: 10.1016/j.bpsc.2019.11.015] [Medline: 32085921]
7. Kalantarian H, Washington P, Schwartz J, Daniels J, Haber N, Wall DP. Guess what?: towards understanding autism from structured video using facial affect. J Healthc Inform Res 2019 Oct;3(1):43-66 [FREE Full text] [doi: 10.1007/s41666-018-0034-9] [Medline: 33313475]
8. Kalantarian H, Jedoui K, Washington P, Wall DP. A mobile game for automatic emotion-labeling of images. IEEE Trans Games 2020 Jun;12(2):213-218 [FREE Full text] [doi: 10.1109/tg.2018.2877325] [Medline: 32551410]
9. Kalantarian H, Washington P, Schwartz J, Daniels J, Haber N, Wall D. A gamified mobile system for crowdsourcing video for autism research. In: 2018 IEEE International Conference on Healthcare Informatics (ICHI).: IEEE; 2018 Presented at: 2018 IEEE International Conference on Healthcare Informatics (ICHI); June 4-7, 2018; New York, NY, United States p. 350-352. [doi: 10.1109/ICHI.2018.00052]
10. Kalantarian H, Jedoui K, Dunlap K, Schwartz J, Washington P, Husic A, et al. The performance of emotion classifiers for children with parent-reported autism: quantitative feasibility study. JMIR Ment Health 2020 Apr;7(4):e13174 [FREE Full text] [doi: 10.2196/13174] [Medline: 32234701]
11. Kalantarian H, Jedoui K, Washington P, Tariq Q, Dunlap K, Schwartz J, et al. Labeling images with facial emotion and the potential for pediatric healthcare. Artif Intell Med 2019 Jul;98:77-86 [FREE Full text] [doi: 10.1016/j.artmed.2019.06.004] [Medline: 31521254]
12. Washington P, Kalantarian H, Kent J, Husic A, Kline A, Leblanc E, et al. Training an emotion detection classifier using frames from a mobile therapeutic game for children with developmental disorders. ArXiv Preprint posted online on Dec 16, 2021. [FREE Full text] [doi: 10.2196/26760]
13. Penev Y, Dunlap K, Husic A, Hou C, Washington P, Leblanc E, et al. A mobile game platform for improving social communication in children with autism: a feasibility study. Appl Clin Inform 2021 Oct;12(5):1030-1040 [FREE Full text] [doi: 10.1055/s-0041-1736626] [Medline: 34788890]
14. Ning M, Daniels J, Schwartz J, Dunlap K, Washington P, Kalantarian H, et al. Identification and quantification of gaps in access to autism resources in the United States: an infodemiological study. J Med Internet Res 2019 Jul;21(7):e13094 [FREE Full text] [doi: 10.2196/13094] [Medline: 31293243]
15. Leblanc E, Washington P, Varma M, Dunlap K, Penev Y, Kline A, et al. Feature replacement methods enable reliable home video analysis for machine learning detection of autism. Sci Rep 2020 Dec;10(1):21245 [FREE Full text] [doi: 10.1038/s41598-020-76874-w] [Medline: 33277527]
16. Washington P, Leblanc E, Dunlap K, Penev Y, Kline A, Paskov K, et al. Precision telemedicine through crowdsourced machine learning: testing variability of crowd workers for video-based autism feature recognition. J Pers Med 2020 Aug;10(3):86 [FREE Full text] [doi: 10.3390/jpm10030086] [Medline: 32823538]
17. Tariq Q, Fleming SL, Schwartz JN, Dunlap K, Corbin C, Washington P, et al. Detecting developmental delay and autism through machine learning models using home videos of Bangladeshi children: development and validation study. J Med Internet Res 2019 Apr;21(4):e13822 [FREE Full text] [doi: 10.2196/13822] [Medline: 31017583]
18. Tariq Q, Daniels J, Schwartz JN, Washington P, Kalantarian H, Wall DP. Mobile detection of autism through machine learning on home video: a development and prospective validation study. PLoS Med 2018 Nov;15(11):e1002705 [FREE Full text] [doi: 10.1371/journal.pmed.1002705] [Medline: 30481180]
19. Abbas H, Garberson F, Glover E, Wall DP. Machine learning approach for early detection of autism by combining questionnaire and home video screening. J Am Med Inform Assoc 2018 Aug;25(8):1000-1007 [FREE Full text] [doi: 10.1093/jamia/ocy039] [Medline: 29741630]
20. Fusaro VA, Daniels J, Duda M, DeLuca TF, D'Angelo O, Tamburello J, et al. The potential of accelerating early detection of autism through content analysis of YouTube videos. PLoS One 2014 Apr;9(4):e93533 [FREE Full text] [doi: 10.1371/journal.pone.0093533] [Medline: 24740236]
21. Chorianopoulou A, Tzinis E, Iosif E, Papoulidi A, Papailiou C, Potamianos A. Engagement detection for children with autism spectrum disorder. In: 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).: IEEE; 2017 Presented at: 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); March 5-9, 2017; New Orleans, LA, USA p. 5055-5059. [doi: 10.1109/icassp.2017.7953119]
22. Rudovic O, Utsumi Y, Lee J, Hernandez J, Ferrer E, Schuller B, et al. CultureNet: a deep learning approach for engagement intensity estimation from face images of children with autism. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).: IEEE; 2018 Presented at: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); October 1-5, 2018; Madrid, Spain p. 339-346. [doi: 10.1109/iros.2018.8594177]
23. Washington P, Tariq Q, Leblanc E, Chrisman B, Dunlap K, Kline A, et al. Crowdsourced privacy-preserved feature tagging of short home videos for machine learning ASD detection. Sci Rep 2021 Apr;11(1):7620 [FREE Full text] [doi: 10.1038/s41598-021-87059-4] [Medline: 33828118]
24. Washington P, Leblanc E, Dunlap K, Penev Y, Varma M, Jung JY, et al. Selection of trustworthy crowd workers for telemedical diagnosis of pediatric autism spectrum disorder. Pac Symp Biocomput 2021;26:14-25 [FREE Full text] [doi: 10.1142/9789811232701_0002] [Medline: 33691000]

25. von Hofsten C, Uhlig H, Adell M, Kochukhova O. How children with autism look at events. Res Autism Spectr Disord 2009 Apr;3(2):556-569. [doi: 10.1016/j.rasd.2008.12.003]
26. Elsabbagh M, Mercure E, Hudry K, Chandler S, Pasco G, Charman T, BASIS Team. Infant neural sensitivity to dynamic eye gaze is associated with later emerging autism. Curr Biol 2012 Feb;22(4):338-342 [FREE Full text] [doi: 10.1016/j.cub.2011.12.056] [Medline: 22285033]
27. Pierce K, Conant D, Hazin R, Stoner R, Desmond J. Preference for geometric patterns early in life as a risk factor for autism. Arch Gen Psychiatry 2011 Jan;68(1):101-109 [FREE Full text] [doi: 10.1001/archgenpsychiatry.2010.113] [Medline: 20819977]
28. Yi L, Feng C, Quinn PC, Ding H, Li J, Liu Y, et al. Do individuals with and without autism spectrum disorder scan faces differently? A new multi-method look at an existing controversy. Autism Res 2014 Feb;7(1):72-83. [doi: 10.1002/aur.1340] [Medline: 24124133]
29. Pusiol G, Esteva A, Hall S, Frank M, Milstein A, Fei-Fei L. Vision-based classification of developmental disorders using eye-movements. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016. 2016 Presented at: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016; October 17-21, 2016; Athens, Greece p. 317-325. [doi: 10.1007/978-3-319-46723-8_37]
30. Riby D, Hancock PJB. Looking at movies and cartoons: eye-tracking evidence from Williams syndrome and autism. J Intellect Disabil Res 2009 Feb;53(2):169-181. [doi: 10.1111/j.1365-2788.2008.01142.x] [Medline: 19192099]
31. Noris B, Barker M, Nadel J, Hentsch F, Ansermet F, Billard A. Measuring gaze of children with Autism spectrum disorders in naturalistic interactions. In: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2011 Presented at: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society; August 30-September 3, 2011; Boston, MA, USA p. 5356-5359. [doi: 10.1109/IEMBS.2011.6091325]
32. Hashemi J, Tepper M, Vallin Spina T, Esler A, Morellas V, Papanikolopoulos N, et al. Computer vision tools for low-cost and noninvasive measurement of autism-related behaviors in infants. Autism Res Treat 2014 Jun;2014:935686 [FREE Full text] [doi: 10.1155/2014/935686] [Medline: 25045536]
33. Chang Z, Di Martino JM, Aiello R, Baker J, Carpenter K, Compton S, et al. Computational methods to measure patterns of gaze in toddlers with autism spectrum disorder. JAMA Pediatr 2021 Aug;175(8):827-836. [doi: 10.1001/jamapediatrics.2021.0530] [Medline: 33900383]
34. Egger HL, Dawson G, Hashemi J, Carpenter KLH, Espinosa S, Campbell K, et al. Automatic emotion and attention analysis of young children at home: a ResearchKit autism feasibility study. NPJ Digit Med 2018 Jun;1:20 [FREE Full text] [doi: 10.1038/s41746-018-0024-6] [Medline: 31304303]
35. Wood E, Baltrusaitis T, Zhang X, Sugano Y, Robinson P, Bulling A. Rendering of eyes for eye-shape registration and gaze estimation. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV). 2015 Presented at: International Conference on Computer Vision (ICCV); December 7-13, 2015; Santiago, Chile p. 3756-3764. [doi: 10.1109/iccv.2015.428]
36. Baltrusaitis T, Zadeh A, Lim Y, Morency L. OpenFace 2.0: facial behavior analysis toolkit. In: 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018).: IEEE; 2018 Presented at: 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018); May 15-19, 2018; Xi'an, China p. 59-66. [doi: 10.1109/fg.2018.00019]
37. Jones W, Klin A. Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. Nature 2013 Dec;504(7480):427-431 [FREE Full text] [doi: 10.1038/nature12715] [Medline: 24196715]
38. Schauder KB, Park WJ, Tsank Y, Eckstein MP, Tadin D, Bennetto L. Initial eye gaze to faces and its functional consequence on face identification abilities in autism spectrum disorder. J Neurodevelop Disord 2019 Dec;11(1):42 [FREE Full text] [doi: 10.1186/s11689-019-9303-z] [Medline: 31883518]
39. Neumann D, Spezio ML, Piven J, Adolphs R. Looking you in the mouth: abnormal gaze in autism resulting from impaired top-down modulation of visual attention. Soc Cogn Affect Neurosci 2006 Dec;1(3):194-202 [FREE Full text] [doi: 10.1093/scan/nsl030] [Medline: 18985106]
40. Pelphrey KA, Sasson NJ, Reznick JS, Paul G, Goldman BD, Piven J. Visual scanning of faces in autism. J Autism Dev Disord 2002 Aug;32(4):249-261. [doi: 10.1023/a:1016374617369] [Medline: 12199131]
41. Chawarska K, Shic F. Looking but not seeing: atypical visual scanning and recognition of faces in 2 and 4-year-old children with autism spectrum disorder. J Autism Dev Disord 2009 Dec;39(12):1663-1672 [FREE Full text] [doi: 10.1007/s10803-009-0803-7] [Medline: 19590943]
42. Guillon Q, Afzali MH, Rogé B, Baduel S, Kruck J, Hadjikhani N. The importance of networking in autism gaze analysis. PLoS One 2015 Oct;10(10):e0141191 [FREE Full text] [doi: 10.1371/journal.pone.0141191] [Medline: 26496498]
43. Guo C, Ashrafian H, Ghafur S, Fontana G, Gardner C, Prime M. Challenges for the evaluation of digital health solutions—a call for innovative evidence generation approaches. NPJ Digit Med 2020 Aug;3(1):110 [FREE Full text] [doi: 10.1038/s41746-020-00314-2] [Medline: 32904379]
44. Chew AMK, Ong R, Lei H, Rajendram M, Verma SK, Fung DSS, et al. Digital health solutions for mental health disorders during COVID-19. Front Psychiatry 2020 Sep;11:582007 [FREE Full text] [doi: 10.3389/fpsyt.2020.582007] [Medline: 33033487]

45. Beede E, Baylor E, Hersch F, Iurchenko A, Wilcox L, Ruamviboonsuk P, et al. A human-centered evaluation of a deep learning system deployed in clinics for the detection of diabetic retinopathy. In: CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.: ACM Digital Library; 2020 Presented at: 2020 CHI Conference on Human Factors in Computing Systems; April 25-30, 2020; Honolulu. [doi: 10.1145/3313831.3376718]
46. Gunasekeran DV, Liu Z, Tan WJ, Koh J, Cheong CP, Tan LH, et al. Evaluating safety and efficacy of follow-up for patients with abdominal pain using video consultation (SAVED study): randomized controlled trial. J Med Internet Res 2020 Jun;22(6):e17417 [FREE Full text] [doi: 10.2196/17417] [Medline: 32459637]
47. Richards D, Enrique A, Eilert N, Franklin M, Palacios J, Duffy D, et al. A pragmatic randomized waitlist-controlled effectiveness and cost-effectiveness trial of digital interventions for depression and anxiety. npj Digit Med 2020 Jun;3:85. [doi: 10.1038/s41746-020-0293-8]

Abbreviations

AOI: area of interest
ASD: autism spectrum disorder
LSTM: long short-term memory
NT: neurotypical

Edited by A Mavragani; submitted 06.07.21; peer-reviewed by D Gunasekeran; comments to author 30.11.21; revised version received 20.12.21; accepted 22.12.21; published 15.02.22.

Please cite as:
Varma M, Washington P, Chrisman B, Kline A, Leblanc E, Paskov K, Stockham N, Jung JY, Sun MW, Wall DP
Identification of Social Engagement Indicators Associated With Autism Spectrum Disorder Using a Game-Based Mobile App: Comparative Study of Gaze Fixation and Visual Scanning Methods
J Med Internet Res 2022;24(2):e31830
URL: https://www.jmir.org/2022/2/e31830
doi:10.2196/31830
PMID:35166683

Original Paper

Effectiveness of Digital Forced-Choice Nudges for Voluntary Data Donation by Health Self-trackers in Germany: Web-Based Experiment

Katharina Pilgrim[1], PhD; Sabine Bohnet-Joschko[1], Prof Dr

Department of Management and Entrepreneurship, Faculty of Management, Economics and Society, Witten Herdecke University, Witten, Germany

Corresponding Author:
Katharina Pilgrim, PhD
Department of Management and Entrepreneurship
Faculty of Management, Economics and Society
Witten Herdecke University
Alfred-Herrhausen-Str 50
Witten, 58455
Germany
Phone: 49 2302926475
Email: katharina.pilgrim@uni-wh.de

Abstract

Background: Health self-tracking is an evidence-based approach to optimize health and well-being for personal self-improvement through lifestyle changes. At the same time, user-generated health-related data can be of particular value for (health care) research. As longitudinal data, these data can provide evidence for developing better and new medications, diagnosing rare diseases faster, or treating chronic diseases.

Objective: This quantitative study aims to investigate the impact of digital forced-choice nudges on the willingness of German health self-trackers to donate self-tracked health-related data for research. This study contributes to the body of knowledge on the effectiveness of nonmonetary incentives. Our study enables a gender-specific statement on influencing factors on the voluntary donation of personal health data and, at the same time, on the effectiveness of digital forced-choice nudges within tracking apps.

Methods: We implemented a digital experiment using a web-based questionnaire by graphical manipulation of the Runtastic tracking app interface. We asked 5 groups independently to indicate their willingness to donate tracked data for research. We used a digital forced-choice nudge via a pop-up window, which framed the data donation request with 4 different counter values. We generated the counter values according to the specific target group needs identified from the research literature.

Results: A sample of 919 was generated, of which, 625 (68%) were women and 294 (32%) were men. By dividing the sample into male and female participants, we take into account research on gender differences in privacy tendencies on the web and offline, showing that female participants display higher privacy concerns than male participants. A statistical group comparison shows that with a small effect size (r=0.21), men are significantly more likely (P=.04) to donate their self-tracked data for research if the need to take on social responsibility is addressed (the prosocial counter value in this case—contributing to society) compared with the control group without counter value. Selfish or pseudoprosocial counter values had no significant effect on willingness to donate health data among male or female health self-trackers in Germany when presented as a forced-choice nudge within a tracking app.

Conclusions: Although surveys regularly reveal an 80% to 95% willingness to donate data on average in the population, our results show that only 41% (377/919) of the health self-trackers would donate their self-collected health data to research. Although selfish motives do not significantly influence willingness to donate, linking data donation to added societal value could significantly increase the likelihood of donating among male self-trackers by 15.5%. Thus, addressing the need to contribute to society promotes the willingness to donate data among male health self-trackers. The implementation of forced-choice framing nudges within tracking apps presented in a pop-up window can add to the accessibility of user-generated health-related data for research.

(J Med Internet Res 2022;24(2):e31363) doi:10.2196/31363

KEYWORDS

quantified self; health self-tracking; digital nudge; data donation; health data; mobile phone

Introduction

Self-tracked Health Data for Research

As of April 2020, more than half a million enthusiastic health self-trackers donated collected data to the federal government research institute responsible for disease control and prevention in Germany, via fitness wristbands or smartwatches [1]. As of February 2021, data records donated have exceeded 170 million sets, including pulse, blood pressure, weight, temperature, physical activity, and information on sleep cycles. Regarding the COVID-19 pandemic, heart rate measurement and physical activity are of particular interest because an accelerated pulse while decreasing activity is a likely fever indicator—a typical COVID-19 symptom. At best, this fever monitor will serve as an early warning system and predict the spread of coronavirus in Germany even before case numbers from public health authorities are available [2].

The approach described is only one instance of how user-generated health-related data could add value to research and support public health measures [3]. Moreover, this kind of real-world data can support the development of pharmaceutical innovation, accelerate rare disease diagnosis, and improve chronic disease treatment [4-10]. For the self-tracked data to be available for research, users have to share, more precisely, donate their data actively to research institutions [11,12]. In addition to observing regulatory, ethical, and privacy issues, it is crucial to know, understand, and evaluate the motives of a specific target group for data donation [13,14]. Research on data donation typically uses a population cross-section sample. Surveys indicate that 4 of 5 Germans are willing to donate their digital health data anonymously and free of charge for medical research [15]. As many as 95% of German and US social media users would specifically donate their data to scientists at universities and public research institutions [15,16]. The results can potentially be highly biased, as individuals may hypothetically signal their willingness to donate data and perhaps not even engage in health self-tracking at all. Thus, on the one hand, research lacks information on the willingness to donate self-tracked health-related data among those who actually engage in health self-tracking for personal reasons and with individual goals. On the other hand, we need to determine whether there are nonmonetary benefits for that target group that can possibly influence the willingness to donate data positively.

Research Question and Objectives

We argue that digital forced-choice nudging in general is an appropriate tool to introduce behavioral change toward donating personal health-related data among health self-trackers because health self-trackers are very likely to share their self-tracked data with third parties and primarily have other intentions for tracking. Thus, data donation for research might be perceived as rather unimportant and therefore present an appropriate field for effective nudging.

This quantitative study aims to investigate whether a digital forced-choice nudge can influence the willingness of German health self-trackers to donate self-tracked health-related data for research. We want to contribute to the body of knowledge on the effectiveness of nonmonetary incentives derived from the known needs of the target population. For this, we also consider research on gender differences in privacy tendencies on the web and offline. This shows that women display higher privacy concerns than men. This is in line with both evolutionary and social role theories. Therefore, we analyzed male and female health self-trackers separately [17,18]. Our study thus enables a gender-specific statement on influencing factors on the voluntary donation of personal health data and, at the same time, on the effectiveness of digital nudges for health self-trackers in Germany.

Background and Theory

Nudging

When discussing the active decision to donate self-collected health data and possible contributing factors, we already know that various psychological effects influence individuals, consciously or unconsciously, during their decision-making process [19,20]. People often act impulsively, emotionally, or simply out of habit [21]; they are not always able to calculate the expected consequences of given options and therefore choose the seemingly best available one [22].

In this context, nudges are tools for influencing behavior in decision-making processes without resorting to prohibition, commandments, or economic incentives [19,23,24]. Thus, a nudge is a nonregulatory approach that attempts to motivate individual behavior change through subtle alterations in the choice environments that people face [19,25], whereas a suggested benefit is embedded in the decision-making process [26]. Thus, Karlsen and Andersen [27] define nudging as a term for influencing decisions and behavior using suggestions, positive reinforcement, and other noncoercive means. Löfgren and Nordblom [28] show that the likelihood of a nudge having an effect is higher for choices that the individual perceives as rather unimportant and at a moment in time with limited attention. In summary, the behavior of our study group, health self-trackers in Germany, can potentially be influenced by a nudge that promises a benefit based on existing target group needs and is implemented at a point of limited attention.

Nudges in Health Systems

Nudges are widely applied in health systems, with the ultimate goal of a healthy population. According to Holland et al [29], most nudging measures aim at a healthier diet, more exercise, and the reduction of alcohol and tobacco consumption and are thus frequently applied in preventive interventions. Especially in the nutrition field, we find a body of evidence for the effectiveness of nudging interventions: the use of nudges such as food traffic lights, the prominent placement of healthy food alternatives, or the transparent display of calories increases the choice of the healthier option by an average of 16.3% in test participants [30]. Okeke et al [31] investigated the impact of haptic (digital detox) nudging (phone vibration) to reduce the time spent on the web to improve users' well-being, successfully reducing daily screen time by over 20%. The growing body of evidence on nudging is also increasingly attracting the attention of health care insurance providers. They can potentially realize considerable savings by encouraging their insureds toward

targeted (healthier) behavior changes while offering the prospect of bonuses in return.

Digital Nudges

With the extensive use of digital devices, decision-making in digital choice environments (so-called digital nudging) has emerged, especially in the field of e-commerce [27,32]. Digital nudges affect value cocreation by (1) widening resource accessibility, (2) extending engagement, or (3) augmenting human actors' agency [33]. Similar to offline nudges, ethical considerations for digital nudges are already discussed in the literature, focusing on topics such as preserving individuals' freedom of choice or autonomy, transparent disclosure of nudges, and individual (proself) and societal (prosocial) goal-oriented justification of nudging [34]. Mirsch et al [32] pointed out that digital nudging in general is a promising research area with great potential for improvement and opportunities, especially regarding user interface, user experience, and digital service design questions. The future potential is supported by findings from the study by Hummel and Maedche [35], who discovered that today, only 62% of digital nudging treatments are statistically significant. Regarding the effect size of digital nudges, a quantitative review showed no difference (median effect size of 21% depending on category and context) compared with offline settings while offering new perspectives of individualization [35]. However, research on smart digital nudges in the health care or data donation field is limited. Meske et al [36], for example, conclude that digital nudging in hospitals can positively influence the use of technology, new value creation, changes in structures, and consequently financial dimensions of digital transformation, supporting caregivers as well as caretakers [36]. Regarding charity program participation, forced-choice nudges are found to be the most efficient ones [37,38]; a review reveals that overall default nudges are most effective, and precommitment strategies are least effective [35].

Hypotheses Development

Overview

To test the effectiveness of digital forced-choice nudges for data donation among health self-trackers in Germany, we need to identify the needs of our test group to use the right triggers for behavior change—in our experiment, donate tracked data. The prospect of need satisfaction could be a potentially attractive reward, which might encourage self-trackers to donate their data in return. To this end, we will derive potentially attractive counter values (benefits) for framing in a digital forced-choice nudge based on the known prevailing motives and needs of our target group and for data donation in general.

Need of Achievement and Power by Self-expertization

People with a diagnosed disease predominantly track vital signs or biological parameters [39-42]. Intrinsic motivation to improve disease management by advancing personal disease knowledge and controlling health indicators (such as glucose or blood pressure levels) is key for tracking [39-42]. Goals are controlling symptoms and preventing or delaying disease progression.

People with self-perceived disease risk factors mainly track their dietary and physical activities. Motives include potential and subjectively perceived risk prevention (such as obesity) or the desire to learn and promote a healthier lifestyle [43,44]. People without a diagnosed or self-perceived disease or prevalence primarily track their nutrition and exercise parameters. Self-design by performance optimization and monitoring performance progress improvement is the motivation behind [45-47]. A less relevant motive is self-entertainment, which includes natural curiosity, a basic interest, and fun in data collection and visualization [46,48,49].

Health self-trackers thus collectively possess a desire to use digital technologies to optimize health and well-being via self-monitoring [50,51]. Self-motivation, self-discipline, or the desire for performance enhancement are motives found in every user group [52,53] and can be labeled as self-expertization [54]. According to McClelland, these motives arise from the need for achievement and power (over a disease or one's own body) [55-59]. As self-expertise is key in every subpopulation with no regard to personal medical conditions, our first hypothesis is as follows:

- H1A: The prospect of receiving individualized tips to improve one's health has a positive influence on female health self-trackers' willingness to donate personal self-collected health-related data for research.
- H1B: The prospect of receiving individualized tips to improve one's health has a positive influence on male health self-trackers' willingness to donate personal self-collected health-related data for research.

Need of Self-actualization by Contributing to Society

Research on the motives to donate personal data, for example, to charity, is consistent with findings on motives supporting prosocial behavior, such as blood donation [60-63]. Donations can positively impact self-image and sense of self, as the donor receives appreciation and care in return [64]. Thus, individual needs as well as the need for self-actualization are also satisfied in this context [65]. Kalkman et al [14] pointed out that although participants recognized the actual or potential benefits of data donation for research, they expressed concerns about confidentiality and data abuse [14]. Nonetheless, 2 positive influences on the willingness to donate personal data exist: social responsibility or sense of duty is the first and most influential factor [66]. It refers to indirect reciprocity: giving back to the community and expecting the same treatment in return [67,68]. This altruistic motive is predominantly based on perceived empathy—the willingness to help out of compassion [69].

Second, an individual's perception of the significance of their own contributions to the community is crucial. Nudging intervention can realize this by emphasizing benefits, such as potentially accelerating the cure of disease by valid research findings and improved therapeutic interventions [16,70].

Derived from identified positive influences on data donation in general and underlying needs, our second hypothesis is as follows:

- H2A: The prospect of contributing to the community has a positive influence on female health self-trackers' willingness to donate personal self-collected health-related data for research.
- H2B: The prospect of contributing to the community has a positive influence on male health self-trackers' willingness to donate personal self-collected health-related data for research.

Need of Recognition and Social Belonging

According to Gimpel et al [46], motives for sharing self-tracked data, such as diet- and exercise-related or vital and biological parameters, can be referred to as self-association. Cost-benefit trade-offs regarding sharing tracked data are strongly linked to situationally perceived experience [71]. On the basis of the social exchange theory by Homan, users unconsciously or consciously weigh the costs or disadvantages of disclosure against personally perceived advantages [72]. Prospects of satisfying feelings of belonging (to a community) and identification with personalized and individual data outweigh possible disadvantages, for example, receiving personalized advertisements or privacy concerns [46]. Self-association can be traced back to social and individual needs satisfaction according to Maslow and Kruntorad [65], as individuals feel an urge for recognition and belonging and, based on this, possess a desire for esteem and prestige. We derive hypotheses 3 and 4 according to the identified motives for personal health-related data sharing as follows:

- H3A: The prospect of recognition by the peer group (classification of personal tracking results) has a positive influence on female health self-trackers' willingness to donate personal self-collected health-related data for research.
- H3B: The prospect of recognition by the peer group (classification of personal tracking results) has a positive influence on male health self-trackers' willingness to donate personal self-collected health-related data for research.
- H4A: The prospect of belonging to a peer group (via personal donation activity) has a positive influence on female health self-trackers' willingness to donate personal self-collected health-related data for research.
- H4B: The prospect of belonging to a peer group (via personal donation activity) has a positive influence on male health self-trackers' willingness to donate personal self-collected health-related data for research.

Methods

Study and Questionnaire Design

To test hypotheses 1 to 4 using digital forced-choice nudges, we set up a web-based experiment with the questionnaire tool LimeSurvey for health self-trackers in Germany. We generated 5 different mock-ups of the tracking app, Runtastic. We chose the app for its continued popularity, familiarity across all age groups, and duration in the market (since 2009—one of the first apps for health self-tracking) [73-75]. These mock-ups depicted the following situation:

the user has just tracked a run of 5.2 km with Runtastic. This screen in particular is a characteristic of a frequently shared status update on Facebook or Twitter [76]. It represents the sharing of personal physical activity by predominantly recreational athletes via social networks [76]. At this point in the user journey within the app, users only want to see their personal stats (and share them).

On completion, a pop-up window with call to action appeared. The pop-up is a new built-in hurdle before reaching the desired results. The person is in a state of physical exertion, having just completed an intense sports session. Attention and interest in pop-up content could be considered lower at this point [28].

Each pop-up window had a recommendation to the user—donating the tracked data to research, followed by information that motivates and helps him choose the suggested behavior—1 of 4 different nudges (N1 to N4) [27]. Presented nudges refer to hypotheses H1A to H4B. N1 is a framing nudge with egoistic benefits, promising individual tips for data donation based on the need for achievement and power through self-experimentation. N2 promises a contribution to society and is thus our prosocial (framing) nudge based on the need for self-actualization by contributing to society. N3 promises the comparison of one's own results with other users (social norm framing), as the second egoistic nudge, built on the identified need for recognition. N4 promises joining the data-for-science community, indicated by a badge within the app, and is thus a pseudoprosocial nudge that also uses social norm framing (belonging to a community). The counter value is based on the identified needs of social belonging and the desire for prestige. Finally, N0 is the control group, with no offered counter value.

Willingness to donate was queried by assessing the likelihood of donating on an 11-point scale, ranging from 0% to 100% (How likely are you to click *Donate Now*?). Only one of the pop-up windows was randomly included in each questionnaire. In addition to three sociodemographic parameters, *gender*, *age*, and *education*, we added 2 items as inclusion criteria to the questionnaire. We started by querying *devices used for tracking health-related data* (*smartphone*, *smartwatch*, *fitness tracker*, or *none*), allowing multiple answers as well as the *frequency of accessing the tracked data* (*daily*, *weekly*, *monthly*, *fever*, or *never)*.

Recruitment

The recruitment strategy included digital social media channels, such as Facebook, Instagram, LinkedIn, Xing, and Twitter. Facebook groups dedicated to fitness and nutrition topics as well as Instagram stories of female fitness microinfluencers represented key channels. To question active health self-trackers in Germany that engage with their tracked data, the defined exclusion criteria were (1) if participants never used a tracking device and (2) if participants never actively or consciously tracked any health-related data.

Data Processing and Statistical Analysis

After collection, the data preparation included cleaning and organizing the raw data set in Excel (Microsoft Inc). We diligently checked for errors to eliminate incomplete questionnaires. Data processing involved encoding text format

data into numeric indicator variables. Ultimately, our sample included exclusively ordinal-scaled variables suitable for statistical analysis in SPSS. We divided the sample into female and male participants to examine gender-specific differences.

To perform appropriate statistical analysis for the experimental evaluation, we checked the data distribution. In addition, we checked for variance homogeneity within the 5 different test groups for each gender. To validate experimental hypotheses H1 to H4, we compared our 5 groups with each other for men and women separately. Regarding the data set characteristics, we applied the Kruskal-Wallis and Mann-Whitney *U* tests as nonparametric tests to compare independent samples with homogeneous variances. The Kruskal-Wallis test can evaluate whether there is an actual effect of group affiliation in the first step. With a subsequent post hoc test, we checked which of the groups differed significantly. We used the Dunn-Bonferroni test.

Results

Sample

We collected 1091 questionnaires in January and February 2021. Following our defined exclusion criteria, we excluded 5.04% (55/1091) of observations because of participants not using a tracking device and 0.73% (8/1091) of observations because they did not track any health-related data actively or consciously. Furthermore, we removed 9.99% (109/1091) of incomplete questionnaires. The final sample consisted of 919 participants.

Our data set was not normally distributed but left-sided (skewness=0.087) and compressed (kurtosis=−1.431; Table 1).

The sample included 68% (625/919) women and 32% (294/919) men. Overall, 45.1% (414/919) of the participants were aged between 18 and 34 years, 46% (423/919) of the participants were aged between 35 and 54 years, and 8.9% (82/919) of the participants were aged >55 years. Overall, only 0.2% (2/919) of participants did not finish high school. Overall, 4.5% (41/919) of the participants were still in school or high school graduates with no additional formal education. In addition, 43.5% (400/919) of the participants were attending or had graduated from college, and 51.8% (476/919) of the participants were going to or had graduated from a university. In terms of tracking frequency, 85.4% (785/919) of the participants reported daily tracking, 10.8% (99/919) of the participants tracked weekly, 1.4% (13/919) of the participants tracked monthly, and 2.4% (22/919) of the participants tracked less than once per month.

We found that 60.3% (554/919) of our sample used a smartwatch for health self-tracking. Overall, 43.2% (397/919) of the participants used a smartphone, and 33% (303/919) of the participants used a fitness tracker (multiple answers were possible).

The sample (N=919) was divided into 5 test groups with N0 as 186 (20.2%; control group without nudge) and 4 different nudge groups (N1=183, 19.9%; N2=163, 17.7%; N3=199, 21.7%; and N4=188, 20.5%; women and men combined).

Taking gender differences in privacy concerns into account, we split the sample into female (Table 2) and male (Table 3) participants for further analysis.

Table 1. Data distribution.

Probability to donate (men + women)	Values
Participants, n (%)	
Valid	919 (100)
Missing	0 (0)
Value, skewness (SE)	0.087 (0.081)
Value, kurtosis (SE)	−1.431 (0.161)

Table 2. Description of the female sample—divided into the 5 test groups (n=625).[a]

Nudge	Participants, n (%)	Value, mean (SD; SE; range; 95% CI)
0	127 (20.3)	41.18 (36.113; 3.205; 0-100; 34.84-47.52)
1	115 (18.4)	43.91 (31.336; 2.922; 0-100; 38.12-49.70)
2	118 (18.9)	47.97 (34.901; 3.213; 0-100; 41.60-54.33)
3	149 (23.8)	45.30 (34.574; 2.832; 0-100; 39.70-50.90)
4	116 (18.6)	34.83 (36.295; 3.370; 0-100; 28.15-41.50)

[a]Total: mean 42.77, SD 34.879; SE 1.395; range 0-100; 95% CI 40.03-45.51.

Table 3. Description of the male sample—divided into the 5 test groups (n=294).[a]

Nudge	Participants, n (%)	Value, mean (SD; SE; range; 95% CI)
0	59 (20.1)	47.80 (36.863; 4.799; 0-100; 38.19-57.40)
1	68 (23.1)	48.24 (36.608; 4.439; 0-100; 39.37-57.10)
2	45 (15.3)	63.33 (33.439; 4.985; 0-100; 53.29-73.38)
3	50 (17)	56.20 (34.279; 4.848; 0-100; 46.46-65.94)
4	72 (26.3)	41.25 (35.759; 4.214; 0-100; 32.85-49.65)

[a]Total: mean 50.10, SD 36.112; SE 2.106; range 0-100; 95% CI 45.96-54.25.

Outcomes of Web-Based Experiment: Nudging Female Health Self-trackers

The female sample contained 625 questionnaires. Regardless of the nudge applied, 77% (481/625) of women surveyed were willing to donate their tracked data for research (with a probability between 10% and 100%).

The Levene test indicates homogeneous variances for female health self-trackers (Table 4).

Using the Kruskal-Wallis test, we examined whether the probability of data donation is the same across the 5 sample groups. As a result, we had to reject the nil hypothesis because significant differences (P=.03) exist across at least 2 groups.

As a post hoc test, we applied the Dunn–Bonferroni test for pairwise group comparison to identify the groups with significant differences. Ultimately, we found no significant difference in the likelihood of donating data between the control group 0 and the 4 test groups. Accordingly, we had to reject the hypotheses H1A, H2A, H3A, and H4A for female health self-trackers, as none of the nudges exerted a significant positive influence on the willingness to donate data. However, there was a significant difference between group 2, the social nudge, and group 4, the prosocial nudge (P=.03; Figure 1; Table 5).

Table 4. Levene test of homogeneity of variances (women; N=625).

Parameters	Levene statistic (df)	Significance (P value)
Probability to donate (women)		
On the basis of the mean	2.007 (4,620)	.09
On the basis of the median	1.323 (4,620)	.26
On the basis of the median and with adjusted df	1.323 (4,548.352)	.26
On the basis of the trimmed mean	1.864 (4,620)	.12

Figure 1. Boxplots for willingness to donate among the 5 test groups (women).

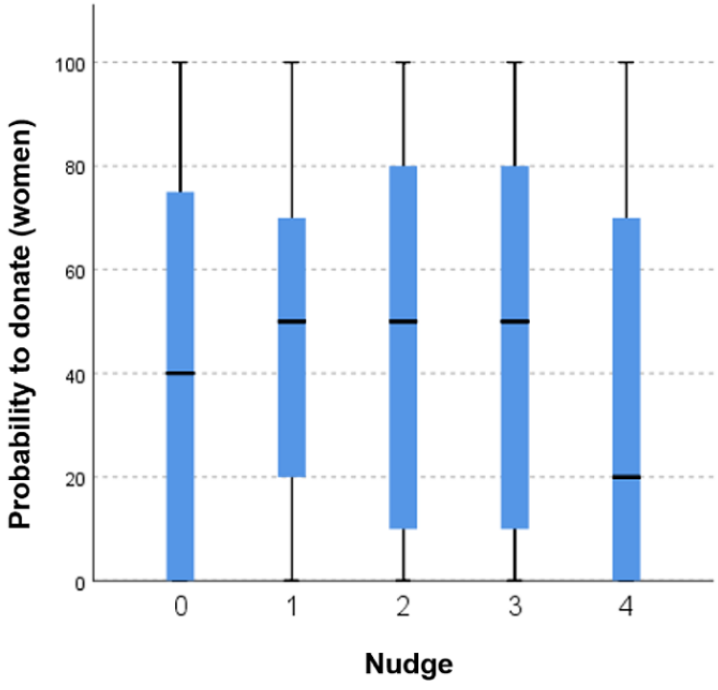


Table 5. Pairwise comparison of nudges (women; N=625).

Sample 1-sample 2[a]	Participants, n (%)	Test statistic (SE)	Standard test statistic	Significance (*P* value)	Adjusted significance[b] (*P* value)
4-0	243 (38.9)	33.037 (22.986)	1.437	.15	.99
4-1	231 (36.9)	52.646 (23.552)	2.235	.03	.25
4-3	265 (42.4)	57.423 (22.161)	2.591	.01	.09
4-2	234 (37.4)	69.431 (23.401)	2.967	.003	.03
0-1	242 (38.7)	−19.609 (23.038)	−.851	.40	.99
0-3	276 (44.2)	−24.386 (21.615)	−1.128	.26	.99
0-2	245 (39.2)	−36.394 (22.884)	−1.590	.11	.99
1-3	264 (42.2)	−4.777 (22.215)	−.215	.83	.99
1-2	233 (37.3)	−16.785 (23.452)	−.716	.47	.99
3-2	267 (42.7)	12.008 (22.055)	.544	.59	.99

[a]Each row tests the null hypothesis that the sample 1 and sample 2 distributions are the same. Asymptotic significances (2-sided tests) are displayed. The significance level is *P*=.05.

[b]Significance values have been adjusted by the Bonferroni correction for multiple tests.

The post hoc test results and plot visualization suggest a potential significant difference between groups 1 and 4 and between groups 3 and 2. We further applied a median-based pairwise group comparison using the Mann–Whitney *U* test. The results indicate significant differences between groups 1 and 4 (N=231; *P*=.01) and between groups 3 and 4 (N=265; *P*=.01).

In summary, for female health self-trackers, no offered counter value, based on egoistic, social, or prosocial motives or needs, has a significant positive or negative influence on the willingness to donate tracked data for research. However, the prospect of showing one's donation behavior to other users has a negative effect on willingness to donate data in direct comparison with groups that receive a prosocial or a self-serving return for their data donation.

Outcomes of Web-Based Experiment: Nudging Male Health Self-trackers

Overall, 79.9% (235/294) of the male respondents would donate their data (with a probability between 10% and 100%), regardless of the nudge queried (3% more than female respondents).

The Levene test indicates homogeneous variances for male health self-trackers (Table 6).

Testing for differences among groups in terms of donation probability also revealed a significant difference among male participants (*P*=.02) using the Kruskal-Wallis test.

A pairwise group comparison also revealed a significant difference between groups 2 (social nudge) and 4 (prosocial nudge; *P*=.02; Figure 2; Table 7).

The Mann-Whitney *U* test revealed significant differences between groups 1 and 2 (*P*=.04), between groups 0 and 2 (*P*=.04), and between groups 3 and 4 (*P*=.02).

As with female health self-trackers, we had to reject hypotheses H1B, H3B, and H4B for men as well. Thus, the prospect of a self-serving benefit or displaying data donation activity, named pseudoprosocial in our experiment, has no significant positive or negative influence on willingness to donate.

On the other hand, the prospect of making a prosocial contribution significantly influences the likelihood of donating self-tracked (health) data with a small effect size (*r*=0.21; *Z*=−2.087; N=104). Accordingly, we could not reject hypothesis H2B. Comparing means, nudging male health self-trackers in their donating decision-making process with a prosocial nudge would increase the willingness to donate data by 15.5%. At the same time, for male health self-trackers, we found evidence that open data donation significantly reduces the probability of donation compared with a *secret donation* (nudge 2), as well as compared with test groups 1 and 3, receiving a personal benefit. In addition, a group comparison of participants receiving personal tips and participants merely making a prosocial contribution reveals again that prosocial reasons are superior to selfish ones and increase the likelihood of donating data in direct comparisons.

By dividing the sample into male and female participants, we considered gender differences because of the known privacy concern differences between men and women. We also split the sample into different groups regarding age, education, devices used for tracking health-related data, and tracking frequency during analysis. The results show that there are no significant increases or decreases in the willingness to donate self-tracked health-related data when the sample is divided into the stated groups. A correlation analysis did not show any significant positive or negative correlations with the willingness to donate any other variable but gender (*P*=.001).

Table 6. Levene test of homogeneity of variances (men; N=294).

Parameters	Levene statistic (df)	Significance (*P* value)
Probability to donate (men)		
On the basis of the mean	0.873 (4,289)	.48
On the basis of the median	1.127 (4,289)	.34
On the basis of the median and with adjusted df	1.127 (4,275.393)	.34
On the basis of the trimmed mean	0.975 (4,289)	.42

Figure 2. Boxplots for willingness to donate among the 5 test groups (men).

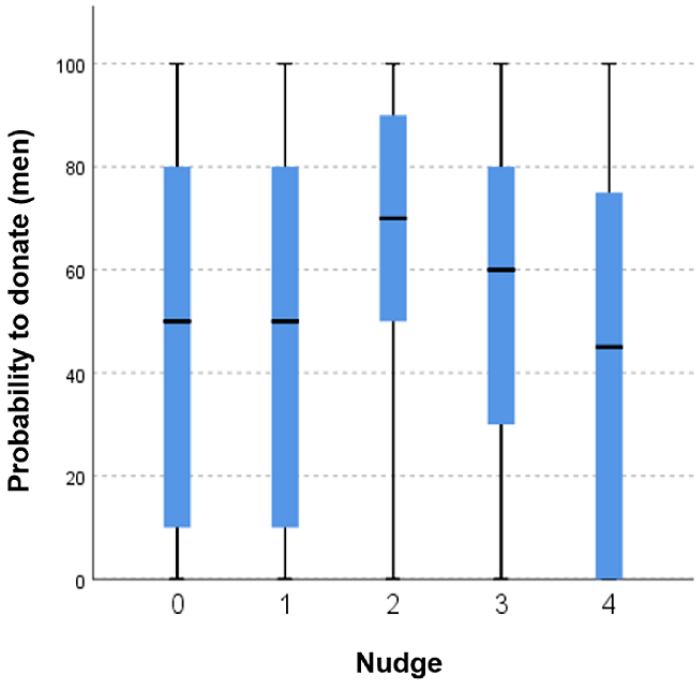


Table 7. Pairwise comparison of nudges (men; N=294).

Sample 1-sample 2[a]	Participants, n (%)	Test statistic (SE)	Standard test statistic	Significance (*P* value)	Adjusted significance[b] (*P* value)
4-0	131 (44.6)	15.441 (14.806)	1.043	.30	.99
4-1	140 (47.6)	17.372 (14.258)	1.218	.22	.99
4-3	122 (41.5)	35.604 (15.522)	2.294	.02	.22
4-2	117 (39.8)	50.646 (16.022)	3.161	.002	.02
0-1	127 (43.2)	−1.932 (15.001)	−.129	.90	.99
0-3	109 (37.1)	−20.163 (16.207)	−1.244	.21	.99
0-2	104 (35.4)	−35.205 (16.687)	−2.110	.04	.35
1-3	118 (40.1)	−18.231 (15.708)	−1.161	.25	.99
1-2	113 (38.4)	−33.274 (16.203)	−2.054	.04	.40
3-2	95 (32.3)	15.042 (17.325)	.868	.39	.99

[a]Each row tests the null hypothesis that the sample 1 and sample 2 distributions are the same. Asymptotic significances (2-sided tests) are displayed. The significance level is *P*=.05.

[b]Significance values have been adjusted by the Bonferroni correction for multiple tests.

Discussion

Principal Findings and Comparison With Previous Work

As apps represent digital environments in which users have to make decisions on an ongoing basis, a neutral choice presentation is impossible from a behavioral economics point of view. User interface design is crucial and influences users' app interactions. Developers need to know and understand user decisions and motives, as well as the inferable effects of design on user decision-making, to support desired actions. Our results show that digital nudges addressing the right user needs seem to be an action-triggering operation. Specifically, a digital forced-choice prosocial framing nudge, presented in a pop-up window, can increase the willingness to donate data among male health self-trackers by 15.5%. The results are in line with research on the effectiveness of digital nudging in general (20%) as well as on the effectiveness of offline nudges in health care (16%) [30,35]. We need to further examine how the nudge itself (framing), the presentation (pop-up window vs no pop-up window), and timing of the digital nudge impact nudge effectiveness. In general, however, developers and researchers should consider digital nudges addressing the need for social responsibility when asking for data donation within this specific user group.

The experimental results confirm the findings of previous research on the motives to donate personal data for the health area. Specifically, for health self-trackers, our study confirms the results of Skatova and Goulding [66] and Mujcic and Leibbrandt [68], who found social responsibility or duty to be the strongest predictor of willingness to donate personal data. Our experiment demonstrated that the prospect of doing something good for society positively influences willingness to donate personal health data. Arguably, the act of donating has a positive impact on self-image and sense of self. Unlike blood donation, which provides immediate reciprocal value to the donor in the form of appreciation and caring (during donation; Schiefer [64]) by those present, our results suggest the opposite for health data donation. The public donation of data and its visibility to third parties, especially peers, can potentially discourage health self-trackers from making this donation—they prefer to do so confidentially. Belonging to an ever more similar group within the group of self-trackers, in this case, people who donate their health data have no positive influence on willingness to donate data and more of a negative influence. Showing *charity*, therefore, negatively influences willingness to donate data. This finding should be taken into account when framing a data donation plea, for example, by explicitly referring to donor anonymity.

Compared with polls regarding willingness to donate data, our results show a significantly lower willingness among active health self-trackers. In contrast to population surveys, which put a willingness to donate data at 80% or social media users' willingness at 95% [15,16], only 41% (377/919) of surveyed health self-trackers would be willing to donate their tracked data to research with a probability above 50% and only 10% (92/919) with 100% probability. This discrepancy suggests that individuals who actively collect health-related data for a particular purpose value it more. It is important to keep this in mind when designing approaches to access health-related, self-tracked data. Research surveying a population cross-section disregards the hypothetical character of questions and answers, and thus, results can be biased and lead to ineffective measures.

Our results also show that women, who are evolutionarily more concerned about protecting their privacy, do not respond to any of the nudges presented adding to the findings by Tifferet [17] and Farinosi and Taipale [18], who found these gender differences, especially in social media users. Regarding the existing gender gap in clinical trials, addressed by Karp and Reavey [77], research needs to investigate which motives, needs, and nudges can increase access to women's health data equally.

Limitations

The findings have a number of limitations. Owing to recruiting primarily via social media and fitness influencers, the sample includes a disproportionately large number of younger and higher educated people as well as mostly female compared with male health self-trackers (2/3 to 1/3), which can bias our results.

Thus, the sample may not represent all German health self-trackers. Furthermore, people engaging in social media and following a call to action from influencers on Instagram and from peers in fitness Facebook groups (participating in the experiment in this case) have fewer privacy concerns and are more likely to share their health-related information with others. We did not perform ex ante power calculations to determine the sample size. Future research may therefore repeat the survey with a larger sample via additional recruitment channels to assess the reliability of our findings. In addition, the chosen geographic focus (Germany) might have biased the results because of cultural differences in terms of relevance and general attitudes toward privacy [78]. Further studies could consider international cross-cultural comparisons to verify the validity of our findings for a global app market.

Participants' nonreproducible or untruthful answers can also limit the results. The reasons could be the chosen and not clearly comprehensible scale levels or phrasing of the individual benefit. Future experiments could also imbed nudges in a different environment (a different app) and use an even more realistic situation with clickable mock-ups. Our hypothetical request for web-based health data donation may not represent reality. Future experiments could implement a real data donation tool within a popular health self-tracking app to verify our results.

Conclusions

The growing trend toward digital health and increasing acceptance as well as the use of health apps such as fitness trackers, digital check-ups, and nutrition apps will contribute to a significant increase in nudging measures. This study could aid access to health data for research and long-term care improvement.

Although selfish motives do not significantly influence willingness to donate, linking data donation to added societal value could significantly increase the likelihood of donating among male self-trackers by 15.5%. Thus, addressing the need

to contribute to society promotes willingness to donate data among male health self-trackers and should be emphasized when designing campaigns to donate health data. The implementation of forced-choice framing nudges within tracking apps presented in a pop-up window can add to the accessibility of user-generated health-related data for research.

Acknowledgments

The study was conducted as part of the ATLAS project, which is funded by the Ministry of Economic Affairs, Innovation, Digitalization, and Energy of North Rhine-Westphalia (funding code: ITG-1-1).

Authors' Contributions

KP conducted the literature research, conducted the quantitative study, and analyzed and discussed the data. SB-J provided content support for the process and reviewed the results. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

References

1. Amft O, Lopera Gonzalez LI, Lukowicz P, Bian S, Burggraf P. Wearables to fight COVID-19: from symptom tracking to contact tracing. IEEE Pervasive Comput 2020;19(4):53-60. [doi: 10.1109/mprv.2020.3021321]
2. Corona data donation app 2.0: donate data and answer surveys - your contribution against Corona. Robert Koch-Institut. 2021. URL: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Corona-Datenspende-allgemein.html [accessed 2020-12-04]
3. Kakkanatt C, Benigno M, Jackson VM, Huang PL, Ng K. Curating and integrating user-generated health data from multiple sources to support healthcare analytics. IBM J Res Dev 2018;62(1):2:1-2:7. [doi: 10.1147/jrd.2017.2756742]
4. Shaw DM, Gross JV, Erren TC. Data donation after death: a proposal to prevent the waste of medical research data. EMBO Rep 2016;17(1):14-17 [FREE Full text] [doi: 10.15252/embr.201541802] [Medline: 26658488]
5. Strotbaum V, Pobiruchin M, Schreiweis B, Wiesner M, Strahwald B. Your data is gold - data donation for better healthcare? Inf Technol 2019;61(5-6):219-229. [doi: 10.1515/itit-2019-0024]
6. Gay V, Leijdekkers P. Bringing health and fitness data together for connected health care: mobile apps as enablers of interoperability. J Med Internet Res 2015;17(11):e260 [FREE Full text] [doi: 10.2196/jmir.5094] [Medline: 26581920]
7. Sharon T. Self-tracking for health and the quantified self: re-articulating autonomy, solidarity, and authenticity in an age of personalized healthcare. Philos Technol 2017;30(1):93-121. [doi: 10.1007/s13347-016-0215-5]
8. Niknejad N, Ismail WB, Mardani A, Liao H, Ghani I. A comprehensive overview of smart wearables: the state of the art literature, recent advances, and future challenges. Eng Appl Artif Intell 2020;90:103529. [doi: 10.1016/j.engappai.2020.103529]
9. Lupton D. Data mattering and self-tracking: what can personal data do? Continuum 2020;34(1):1-13. [doi: 10.1080/10304312.2019.1691149]
10. Lupton D. Self-tracking modes: reflexive self-monitoring and data practices. SSRN J 2014:1-19. [doi: 10.2139/ssrn.2483549]
11. Bietz MJ, Bloss CS, Calvert S, Godino JG, Gregory J, Claffey MP, et al. Opportunities and challenges in the use of personal health data for health research. J Am Med Inform Assoc 2016;23(e1):e42-e48 [FREE Full text] [doi: 10.1093/jamia/ocv118] [Medline: 26335984]
12. Chen J, Bauman A, Allman-Farinelli M. A study to determine the most popular lifestyle smartphone applications and willingness of the public to share their personal data for health research. Telemed J E Health 2016;22(8):655-665. [doi: 10.1089/tmj.2015.0159] [Medline: 26958742]
13. Bietz M, Patrick K, Bloss C. Data donation as a model for citizen science health research. Citiz Sci Theory Pract 2019;4(1):6. [doi: 10.5334/cstp.178]
14. Kalkman S, van Delden J, Banerjee A, Tyl B, Mostert M, van Thiel G. Patients' and public views and attitudes towards the sharing of health data for research: a narrative review of the empirical evidence. J Med Ethics 2022;48(1):3-13 [FREE Full text] [doi: 10.1136/medethics-2019-105651] [Medline: 31719155]
15. Semler SC. "Datenspende“ für die medizinische Forschung: ergebnisse einer aktuellen umfrage. Medizininformatik-initiative. 2019. URL: https://www.tmf-ev.de/DesktopModules/Bring2mind/DMX/Download.aspx?EntryId=32081&PortalId=0 [accessed 2021-01-07]
16. Grajales F, Clifford D, Loupos P, Okun S, Quattrone S, Simon M, et al. Social networking sites and the continuously learning health system: a survey. NAM Perspectives 2014;4(1):1-26. [doi: 10.31478/201401d]
17. Tifferet S. Gender differences in privacy tendencies on social network sites: a meta-analysis. Comput Hum Behav 2019;93:1-12. [doi: 10.1016/j.chb.2018.11.046]

18. Farinosi M, Taipale S. Who can see my stuff? Online self-disclosure and gender differences on Facebook. OBS* 2018;12(1):53-71. [doi: 10.15847/obsobs12120181129]
19. Mirsch T, Lehrer C, Jung R. Digital nudging: altering user behavior in digital environments. In: Proceedings der 13th Internationalen Tagung Wirtschaftsinformatik. 2017 Presented at: WI '17; February 12-15, 2017; St. Gallen.
20. Weinmann M, Schneider C, vom Brocke J. Digital nudging. Bus Inf Syst Eng 2016;58(6):433-436. [doi: 10.2139/ssrn.2708250]
21. Hechter M, Kanazawa S. Sociological rational choice theory. Annu Rev Sociol 1997;23(1):191-214. [doi: 10.1146/annurev.soc.23.1.191]
22. Müller H. Spielen hilft nicht immer. die grenzen des rational-choice-ansatzes und der platz der theorie kommunikativen handelns in der analyse internationaler beziehungen. Zeitschrift für internationale beziehungen 1995;2(2):371-391 [FREE Full text]
23. Thaler RH, Sunstein CR. NUDGE: improving decision about health, wealth, and happiness. London: Penguin; 2009:185-198.
24. Sunstein CR. Nudging: a very short guide. J Consum Policy 2014;37(4):583-588. [doi: 10.1007/s10603-014-9273-1]
25. Oliver A. From nudging to budging: using behavioural economics to inform public sector policy. J Soc Policy 2013;42(4):685-700. [doi: 10.1017/s0047279413000299]
26. Chong D, Druckman JN. Framing theory. Annu Rev Polit Sci 2007;10(1):103-126. [doi: 10.1146/annurev.polisci.10.072805.103054]
27. Karlsen R, Andersen A. Recommendations with a nudge. Technologies 2019;7(2):45. [doi: 10.3390/technologies7020045]
28. Löfgren Å, Nordblom K. A theoretical framework of decision making explaining the mechanisms of nudging. J Econ Behav Organ 2020;174:1-12. [doi: 10.1016/j.jebo.2020.03.021]
29. Hollands GJ, Shemilt I, Marteau TM, Jebb SA, Kelly MP, Nakamura R, et al. Altering micro-environments to change population health behaviour: towards an evidence base for choice architecture interventions. BMC Public Health 2013;13:1218 [FREE Full text] [doi: 10.1186/1471-2458-13-1218] [Medline: 24359583]
30. Arno A, Thomas S. The efficacy of nudge theory strategies in influencing adult dietary behaviour: a systematic review and meta-analysis. BMC Public Health 2016;16:676 [FREE Full text] [doi: 10.1186/s12889-016-3272-x] [Medline: 27475752]
31. Okeke F, Sobolev M, Dell N, Estrin D. Good vibrations: can a digital nudge reduce digital overload? In: Proceedings of the 20th international conference on human-computer interaction with mobile devices and services. New York: Association for Computing Machinery; 2018 Presented at: MobileHCI '18; September 3-6, 2018; Barcelona p. 1-12. [doi: 10.1145/3229434.3229463]
32. Mirsch T, Lehrer C, Jung R. Making digital nudging applicable: the digital nudge design method. In: Proceedings of the 39th international conference on information systems.: Association for Information Systems. AIS Electronic Library (AISeL); 2018 Presented at: ICIS '18; December 13-16, 2018; San Francisco.
33. Mele C, Russo Spena T, Kaartemo V, Marzullo ML. Smart nudging: how cognitive technologies enable choice architectures for value co-creation. J Bus Res 2021;129:949-960. [doi: 10.1016/j.jbusres.2020.09.004]
34. Lembcke TB, Engelbrecht N, Brendel AB, Kolbe L. To nudge or not to nudge: ethical considerations of digital nudging based on its behavioral economics roots. In: Proceedings of 27th European conference on information systems. 2019 Presented at: ECIS '19; June 8-14, 2019; Stockholm.
35. Hummel D, Maedche A. How effective is nudging? A quantitative review on the effect sizes and limits of empirical nudging studies. J Behav Exp Econ 2019;80:47-58. [doi: 10.1016/j.socec.2019.03.005]
36. Meske C, Amojo I, Poncette AS, Balzer F. The potential role of digital nudging in the digital transformation of the healthcare industry. In: 21st HCII: international conference on human-computer interaction. 2019 Presented at: HCII '19; July 26-31, 2019; Orlando p. 323-336. [doi: 10.1007/978-3-030-23538-3_25]
37. Meske C, Amojo I, Mohr P. Digital nudging to increase usage of charity features on e-commerce platforms. In: 15th international conference on business informatics. 2020 Presented at: WI '20; March 9-11, 2020; Potsdam. [doi: 10.30844/wi_2020_k5-meske]
38. Liebig G, Rommel J. Active and forced choice for overcoming status quo bias: a field experiment on the adoption of “no junk mail” stickers in Berlin, Germany. J Consum Policy 2014;37(3):423-435. [doi: 10.1007/s10603-014-9264-2]
39. Karkar R, Zia J, Vilardaga R, Mishra SR, Fogarty J, Munson SA, et al. A framework for self-experimentation in personalized health. J Am Med Inform Assoc 2016;23(3):440-448 [FREE Full text] [doi: 10.1093/jamia/ocv150] [Medline: 26644399]
40. Lee HR, Panont WR, Plattenburg B, de la Croix JP, Patharachalam D, Abowd G. Asthmon: empowering asthmatic children's self-management with a virtual pet. In: CHI'10 extended abstracts on human factors in computing systems. 2010 Presented at: CHI EA '10; April 10-15, 2010; Atlanta p. 3583-3588. [doi: 10.1145/1753846.1754022]
41. Walters DL, Sarela A, Fairfull A, Neighbour K, Cowen C, Stephens B, et al. A mobile phone-based care model for outpatient cardiac rehabilitation: the care assessment platform (CAP). BMC Cardiovasc Disord 2010;10:5 [FREE Full text] [doi: 10.1186/1471-2261-10-5] [Medline: 20109196]
42. Klasnja P, Hartzler A, Powell C, Phan G, Pratt W. Health weaver mobile: designing a mobile tool for managing personal health information during cancer care. AMIA Annu Symp Proc 2010;2010:392-396 [FREE Full text] [Medline: 21347007]

43. Miller AD, Mynatt ED. StepStream: a school-based pervasive social fitness system for everyday adolescent health. In: Proceedings of the SIGCHI conference on human factors in computing systems. New York: Association for Computing Machinery; 2014 Presented at: CHI '14; April 16-May 1, 2014; Toronto p. 2823-2832. [doi: 10.1145/2556288.2557190]
44. Cordeiro F, Epstein DA, Thomaz E, Bales E, Jagannathan AK, Abowd G, et al. Barriers and negative nudges: exploring challenges in food journaling. Proc SIGCHI Conf Hum Factor Comput Syst 2015;2015:1159-1162 [FREE Full text] [doi: 10.1145/2702123.2702155] [Medline: 26894233]
45. Choe EK, Lee NB, Lee B, Pratt W, Kientz JA. Understanding quantified-selfers' practices in collecting and exploring personal data. In: 32nd annual ACM conference on human factors in computing systems. New York: Association for Computing Machinery; 2014 Presented at: CHI '14; April 26-May 1, 2014; Toronto p. 1143-1152. [doi: 10.1145/2556288.2557372]
46. Gimpel H, Nißen M, Görlitz R. Quantifying the quantified self: a study on the motivations of patients to track their own health. In: Proceedings of 34th international conference on information systems. 2013 Presented at: ICIS '13; December 15-18, 2013; Milan.
47. Swan M. Sensor mania! The Internet of things, wearable computing, objective metrics, and the quantified self 2.0. J Sens Actuator Netw 2012;1(3):217-253. [doi: 10.3390/jsan1030217]
48. Li I, Dey AK, Forlizzi J. Understanding my data, myself: supporting self-reflection with ubicomp technologies. In: Proceedings of the 13th international conference on ubiquitous computing. New York: Association for Computing Machinery; 2011 Presented at: UbiComp '11; September 17-21, 2011; Beijing p. 405-414. [doi: 10.1145/2030112.2030166]
49. Budzinski O, Schneider S. Smart fitness: Ökonomische effekte einer digitalisierung der selbstvermessung. List Forum 2017;43:89-124. [doi: 10.1007/s41025-017-0059-5]
50. Lupton D. 'It's made me a lot more aware': a new materialist analysis of health self-tracking. Media Int Aust 2019;171(1):66-79. [doi: 10.1177/1329878x19844042]
51. Fox NJ. Personal health technologies, micropolitics and resistance: a new materialist analysis. Health (London) 2017;21(2):136-153. [doi: 10.1177/1363459315590248] [Medline: 26216896]
52. Sanders R. Self-tracking in the digital era: biopower, patriarchy, and the new biometric body projects. Body Soc 2017;23(1):36-63. [doi: 10.1177/1357034x16660366]
53. Walz SP, Deterding S. Foucault's fitbit: governance and gamification. In: Walz SP, Deterding S, editors. The gameful world: approaches, issues, applications. Cambridge: MIT Press; 2014:339-358.
54. Heyen NB. From self-tracking to self-expertise: the production of self-related knowledge by doing personal science. Public Underst Sci 2020;29(2):124-138 [FREE Full text] [doi: 10.1177/0963662519888757] [Medline: 31778095]
55. McClelland D. Achievement motivation theory. In: Miner JB, editor. Organizational behavior 1: essential theories of motivation and leadership. Milton park: Routledge; 2005:46-60.
56. Fotopoulou A, O'Riordan K. Training to self-care: fitness tracking, biopedagogy and the healthy consumer. Health Sociol Rev 2016;26(1):54-68. [doi: 10.1080/14461242.2016.1184582]
57. Lentferink AJ, Oldenhuis HK, de Groot M, Polstra L, Velthuijsen H, van Gemert-Pijnen JE. Key components in eHealth interventions combining self-tracking and persuasive eCoaching to promote a healthier lifestyle: a scoping review. J Med Internet Res 2017;19(8):e277 [FREE Full text] [doi: 10.2196/jmir.7288] [Medline: 28765103]
58. Salamati F, Pasek ZJ. Personal wellness: complex and elusive product and distributed self-services. Procedia CIRP 2014;16:283-288. [doi: 10.1016/j.procir.2014.02.016]
59. Institute of Medicine (US) Committee on Prevention of Mental Disorders, Mrazek PJ, Haggerty RJ. Reducing risks for mental disorders: frontiers for preventive intervention research. Washington, D.C: National Academies Press; 1994:1-32.
60. Skatova A, Ng E, Goulding J. Data donation: Sharing personal data for public good. In: Digital economy all hands meeting 2014. 2014 Presented at: BESiDE '14; December 3-5, 2014; London. [doi: 10.13140/2.1.2567.8405]
61. Evans R, Ferguson E. Defining and measuring blood donor altruism: a theoretical approach from biology, economics and psychology. Vox Sang 2013;106(2):118-126. [doi: 10.1111/vox.12080] [Medline: 24117697]
62. Penner LA, Dovidio JF, Piliavin JA, Schroeder DA. Prosocial behavior: multilevel perspectives. Annu Rev Psychol 2005;56:365-392. [doi: 10.1146/annurev.psych.56.091103.070141] [Medline: 15709940]
63. de Waal FB. Putting the altruism back into altruism: the evolution of empathy. Annu Rev Psychol 2008;59:279-300. [doi: 10.1146/annurev.psych.59.103006.093625] [Medline: 17550343]
64. Schiefer G. Motive des blutspendens: tiefenpsychologische untersuchung mit gestaltungsoptionen für das marketing von nonprofit-organisationen des blutspendewesens. Cham: Springer; 2007:93-119.
65. Maslow AH, Kruntorad P. Motivation und persönlichkeit. Hamburg: Rowohlt; 1981:49-61.
66. Skatova A, Goulding J. Psychology of personal data donation. PLoS One 2019;14(11):e0224240 [FREE Full text] [doi: 10.1371/journal.pone.0224240] [Medline: 31747408]
67. Godin G, Conner M, Sheeran P, Bélanger-Gravel A, Germain M. Determinants of repeated blood donation among new and experienced blood donors. Transfusion 2007;47(9):1607-1615. [doi: 10.1111/j.1537-2995.2007.01331.x] [Medline: 17725724]
68. Mujcic R, Leibbrandt A. Indirect reciprocity and prosocial behaviour: evidence from a natural field experiment. Econ J 2018;128(611):1683-1699. [doi: 10.1111/ecoj.12474]

69. Titchener EB. Lectures on the experimental psychology of the thought-processes. Basingstoke: Macmillan; 1909:157-194.
70. Hummel P, Braun M, Dabrock P. Data donations as exercises of sovereignty. In: Krutzinna J, Floridi L, editors. The ethics of medical data donation. Cham: Springer; 2019:23-54.
71. Liao Y. Sharing personal health information on social media: balancing self-presentation and privacy. In: Proceedings of the 10th international conference on social media and society. New York: Association for Computing Machinery; 2019 Presented at: SMSociety '19; July 19-21, 2019; Toronto p. 194-204. [doi: 10.1145/3328529.3328560]
72. Krasnova H, Spiekermann S, Koroleva K, Hildebrand T. Online social networks: why we disclose. J Inf Technol 2010;25(2):109-125. [doi: 10.1057/jit.2010.6]
73. Fröhlich G. Die Geschichte medienbasierter Selbsttechnologien von Rousseau bis Runtastic. In: Gentzel P, Krotz F, Wimmer J, Winter R, editors. Das vergessene Subjekt. Cham: Springer; 2019:207-226.
74. Monroy Antón A, Rodríguez Rodríguez B. Runtastic PRO app: an excellent all-rounder for logging fitness. Br J Sports Med 2016;50(11):705-706. [doi: 10.1136/bjsports-2015-094940] [Medline: 26294442]
75. Ittner L, Novak F. Über den Faktor Zeit: Warum Timing alles ist – und es nur selten Sinn macht, ausschließlich auf organisches Wachstum zu setzen. In: Business Angel Hansi Hansmann. Wiesbaden: Springer; 2018:61-66.
76. Stragier J, Evens T, Mechant P. Broadcast yourself: an exploratory study of sharing physical activity on social networking sites. Media Int Aust 2015;155(1):120-129. [doi: 10.1177/1329878x1515500114]
77. Karp NA, Reavey N. Sex bias in preclinical research and an exploration of how to change the status quo. Br J Pharmacol 2019;176(21):4107-4118 [FREE Full text] [doi: 10.1111/bph.14539] [Medline: 30418665]
78. Tsoi HK, Chen L. From privacy concern to uses of social network sites: a cultural comparison via user survey. In: IEEE 3rd international conference on privacy, security, risk and trust and IEEE 3rd international conference on social computing.: IEEE; 2011 Presented at: SocialCom and PASSAT '11; October 9-11, 2011; Boston p. 457-464. [doi: 10.1109/passat/socialcom.2011.71]

Edited by R Kukafka; submitted 21.06.21; peer-reviewed by M Salimi, T Karen, O El-Gayar, J Nafziger; comments to author 22.09.21; revised version received 17.11.21; accepted 17.12.21; published 21.02.22.

Please cite as:
Pilgrim K, Bohnet-Joschko S
Effectiveness of Digital Forced-Choice Nudges for Voluntary Data Donation by Health Self-trackers in Germany: Web-Based Experiment
J Med Internet Res 2022;24(2):e31363
URL: https://www.jmir.org/2022/2/e31363
doi:10.2196/31363
PMID:35188472

Original Paper

The Use of Wearable Pulse Oximeters in the Prompt Detection of Hypoxemia and During Movement: Diagnostic Accuracy Study

Mauro Santos[1,2*], DPhil; Sarah Vollam[2,3*], PhD; Marco AF Pimentel[1], DPhil; Carlos Areia[2,3], MSc; Louise Young[2,3], BSc; Cristian Roman[1,2], PhD; Jody Ede[3], MSc; Philippa Piper[4], BSc; Elizabeth King[4,5], BSc; Mirae Harford[2,3], MD; Akshay Shah[6], DPhil; Owen Gustafson[4], MSc; Lionel Tarassenko[1,2], DPhil; Peter Watkinson[2,3,4], MD

[1]Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, United Kingdom

[2]National Institute for Health Research Oxford Biomedical Research Centre, Oxford, United Kingdom

[3]Critical Care Research Group, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, United Kingdom

[4]Adult Intensive Care Unit, Oxford University Hospitals National Health Service Foundation Trust, Oxford, United Kingdom

[5]Therapies Clinical Service, Oxford University Hospitals National Health Service Foundation Trust, Oxford, United Kingdom

[6]Radcliffe Department of Medicine, University of Oxford, Oxford, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Sarah Vollam, PhD
Critical Care Research Group
Nuffield Department of Clinical Neurosciences
University of Oxford
Kadoorie Centre, Level 3
John Radcliffe Hospital
Oxford, OX3 9DU
United Kingdom
Phone: 44 01865223058
Email: sarah.vollam@ndcn.ox.ac.uk

Abstract

Background: Commercially available wearable (ambulatory) pulse oximeters have been recommended as a method for managing patients at risk of physiological deterioration, such as active patients with COVID-19 disease receiving care in hospital isolation rooms; however, their reliability in usual hospital settings is not known.

Objective: We report the performance of wearable pulse oximeters in a simulated clinical setting when challenged by motion and low levels of arterial blood oxygen saturation (SaO_2).

Methods: The performance of 1 wrist-worn (Wavelet) and 3 finger-worn (CheckMe O2+, AP-20, and WristOx2 3150) wearable, wireless transmission–mode pulse oximeters was evaluated. For this, 7 motion tasks were performed: at rest, sit-to-stand, tapping, rubbing, drinking, turning pages, and using a tablet. Hypoxia exposure followed, in which inspired gases were adjusted to achieve decreasing SaO_2 levels at 100%, 95%, 90%, 87%, 85%, 83%, and 80%. Peripheral oxygen saturation (SpO_2) estimates were compared with simultaneous SaO_2 samples to calculate the root-mean-square error (RMSE). The area under the receiver operating characteristic curve was used to analyze the detection of hypoxemia (ie, SaO_2<90%).

Results: SpO_2 estimates matching 215 SaO_2 samples in both study phases, from 33 participants, were analyzed. Tapping, rubbing, turning pages, and using a tablet degraded SpO_2 estimation (RMSE>4% for at least 1 device). All finger-worn pulse oximeters detected hypoxemia, with an overall sensitivity of ≥0.87 and specificity of ≥0.80, comparable to that of the Philips MX450 pulse oximeter.

Conclusions: The SpO_2 accuracy of wearable finger-worn pulse oximeters was within that required by the International Organization for Standardization guidelines. Performance was degraded by motion, but all pulse oximeters could detect hypoxemia. Our findings support the use of wearable, wireless transmission–mode pulse oximeters to detect the onset of clinical deterioration in hospital settings.

Trial Registration: ISRCTN Registry 61535692; http://www.isrctn.com/ISRCTN61535692

International Registered Report Identifier (IRRID): RR2-10.1136/bmjopen-2019-034404

(J Med Internet Res 2022;24(2):e28890) doi:10.2196/28890

KEYWORDS

diagnostic accuracy; hypoxia; hypoxemia; wearable pulse oximeter; continuous monitoring; mHealth; wearable technology; patient monitoring; deterioration; blood oxygen; hospital

Introduction

Failure to recognize and act on physiological indicators of worsening acute illness in hospital wards is a prevalent problem first recognized over 20 years ago [1-3]. Current practice involves intermittent measurements of vital signs and use of early warning scores [4], which are limited by the intermittent nature of the measurements and the associated time burden for staff [5]. Monitoring vital signs continuously with wearable (ambulatory) devices may overcome these limitations and improve detection of deterioration [6-8]. However, recent pilot and observational studies of wearable monitoring devices have shown mixed results, and no large clinical trials of ambulatory monitoring systems (AMSs) have demonstrated improved patient outcomes [6,9,10]. For example, in pulse oximetry, it is well known that patient motion and low perfusion in extremities can generate artifacts that reduce the accuracy of peripheral oxygen saturation (SpO_2) readings [11]. This represents a major barrier to the deployment of these wearable devices for in-hospital patient monitoring [12]. Data averaging, alarm delay, and data holding are some of the strategies developed by pulse oximeter manufacturers to reduce the effect of motion artifacts and avoid false alerts [13], but there is still a need for studies of diagnostic accuracy and motion artifacts to support development of reliable wearable devices [3,7,14,15]. This need has become acute as health care systems have recommended the incorporation of ambulatory pulse oximeters in the home management of COVID-19 [16-18].

This study is part of a phased mixed-methods research project aiming to develop and refine an AMS using wearable devices to aid in the detection of deterioration and improve patient outcomes. The primary objective of this study was to determine the specificity and sensitivity of currently available ambulatory vital sign–monitoring equipment for the detection of hypoxemia. The secondary objective was to determine the effect of motion on data acquisition by the same devices.

Methods

Ethics

This research publication follows the Standards for Reporting Diagnostic Accuracy Studies reporting guidelines [19] and reports the results of the study protocol of Areia et al [20]. This study received ethics approval from the East of Scotland Research Ethics Service REC 2 (19/ES/0008) and was registered in June 2019 (no. ISRCTN61535692).

Study Design

This was a prospective, observational study in which SpO_2 estimates from the study devices were compared with the gold-standard arterial blood oxygen saturation (SaO_2) samples and clinical-standard SpO_2 estimates collected from arterial blood gas (ABG) samples and a nonambulatory Philips MX450 (Philips, Amsterdam, the Netherlands) pulse oximeter, respectively. The device's pulse rate estimation accuracy is reported in Multimedia Appendix 1.

Participants

Healthy adults (18 years or older) able to give informed consent for participation in the study were recruited consecutively from the Oxford area (United Kingdom) between June 18 and August 8, 2019. The exclusion criteria are described in detail in the study protocol [20], including clinical conditions that might bias the estimation of SpO_2 by oximetry (eg, anemia) or increase risk to the participants' health (eg, clotting disorders).

Test Methods

Study Sessions

The study sessions took place at the Cardiovascular Clinical Research Facility, John Radcliffe Hospital, Oxford, UK. An arterial line was first inserted, under local anesthesia, preferentially into the nondominant radial artery of participants placed in the semirecumbent position (30^o head up). Where it was not possible to cannulate the nondominant arm, the dominant arm was cannulated. Participants wore 1 wrist-only device (Wavelet; WaveletHealth, Mountain View, USA) and 3 wrist-worn pulse oximeters with a finger probe: CheckMe O2+ (Viatom Technology Co Ltd, Shenzhen, China), AP-20 (Shenzhen Creative Industry Co Ltd, Shenzhen, China), and WristOx2 3150 with Bluetooth Low Energy (BLE; Nonin Medical Inc, Plymouth, USA) on the same arm. These devices are among the few that make both numeric and waveform data available to other systems. This is a requirement in our research [20] as we plan to notify clinical staff about the signal quality of the waveforms from which the numeric estimates are derived. A nonambulatory Philips MX450 pulse oximeter was also worn. CheckMe O2+ was always placed on the first finger as per the manufacturer's recommendation. The position of the other 3 finger probes on the second, third, and fourth fingers was randomized using software from Haahr [21], per study visit day, ensuring an even distribution of placement. The participants also wore a 3-lead electrocardiogram (ECG) and an end-tidal carbon dioxide monitor connected to the Philips MX450 monitor, and an adhesive chest patch, for monitoring and acquisition of the heart rate and breathing rate. Results obtained with the chest patch are not reported here.

Stage 1: Movement Phase

An at-rest window was assigned to the period before the first ABG measurement, taken after fitting all the devices. The

participants then moved to a chair and were asked to complete a series of consecutive motion tasks: 20 times sit-to-stand (STS), 2-minute tapping at 2 Hz, 2-minute rubbing at 2 Hz, 20 times drinking from a plastic cup, 50 times turning pages, and a set of predefined tablet activity tasks [20]. ABG measurements were made at the end of each motion task in order to analyze the mean bias of the SpO_2 estimates for each task. For a sample of 15 participants, an additional ABG measurement was made in the middle of the STS motion to assess differences in the SaO_2 during and after that activity.

Stage 2: Hypoxia Exposure Phase

Participants moved to a semirecumbent, supine position and wore a tight-fitting silicone facemask connected to a device that reduces the inspired fraction of oxygen, the hypoxicator unit (Everest Summit Hypoxic Generator, Altitude Centre, London, UK). During this phase, oxygen saturation from the clinical-standard Philips MX450 monitor guided the titration of the hypoxicator by a senior anesthetist from the research team, with appropriate resuscitation facilities nearby. In addition, 7% oxygen in nitrogen was used to further lower the fraction of inspired oxygen (FiO_2), if required [22]. FiO_2 was also monitored via an in-line gas analyzer. ABGs were sampled when the participants reached stable prespecified target peripheral oxygen saturation levels: 95%, 90%, 87%, 85%, 83%, and 80%. A senior anesthetist decided when a stable oxygen level was achieved in order to take the ABG based on the clinical values shown by the standard SpO_2 monitor.

Data Collection

Demographic data, including age, sex, height, weight, skin type (Fitzpatrick scale [23]), baseline heart rate (Philips MX450 3-lead ECG), and SaO_2 (from the initial ABG), were collected for each participant at the start of their session. All data collection devices and software were synchronized to the same timestamp, at the start of each study session. SpO_2 data (1 Hz) from CheckMe O2+ and WristOx2 3150 were sent via BLE to and timestamped in different Android tablets (application developed in-house). The AP-20 SpO_2 data (1 Hz) were captured in the device and then downloaded via Oximeter Data Manager version 5.6 software (Shenzhen Creative Industry Co. Ltd., China). The Wavelet device first uploaded the photoplethysmography data to its web platform via an iOS app "On-site." The platform then retrospectively estimated SpO_2 (1 Hz), and these data were shared with the research team. The Phillips MX450 SpO_2 data were collected using ixTrend version 2.1 software. The start and stop times of each motion task and the ABG measurement timings were recorded in case report forms. Functional SaO_2 values were determined immediately after each ABG sample was taken, by multiwavelength oximetry, using a calibrated blood gas electrolyte analyzer, Radiometer ABL90 Flex (Radiometer, Copenhagen, Denmark).

Statistical Analysis

Sample Size

The sample size calculation was based on the International Organization for Standardization (ISO) 80601-2-61:2019 guideline for testing the accuracy of pulse oximeters, which requires at least 200 data points balanced across the SaO_2 range of 70%-100% from at least 10 subjects. We aimed to collect approximately 30 full data sets (with 7 ABGs being used in both the movement and hypoxia exposure phases, yielding a total of 420 readings, ie, 210 for each phase) to achieve a sufficient number of data points for the primary and secondary outcomes, and to recruit participants varying in their physical characteristics to the greatest extent possible. We excluded participants if incomplete data were collected for any 1 device during testing or if hypoxia was not achieved.

Accuracy, Bias, and Precision Metrics

Demographics and baseline vital sign descriptors were summarized using the mean, the median, and the first and third quartiles for continuous variables and proportions for categorical variables. In accordance with the ISO guideline, the accuracy of the SpO_2 estimates for each device was determined using the root-mean-square error (RMSE) between the measured values (SpO_{2i}) and the reference values (SaO_{2i}):

The RMSE 95% CI was determined using bootstrapping (random sampling with replacement) with 10,000 repetitions. The ISO guideline requires that valid oximeters present an RMSE below or equal to 4% (and below or equal to 8% when considering the CI). To interpret potential sources of the SpO_2 estimation error, the mean bias B and precision S were also calculated as

and

respectively. The latter is also known as the SD of the residuals, which determines the spread of the test SpO_2 data around the linear regression model, SpO_{2fit}, which predicts the SpO_2 estimates that best fit the reference SaO_2 values. The agreement between the test devices and the gold standard was also assessed via Bland-Altman plots. Finally, the mean absolute bias was also analyzed.

Movement Phase

The metrics were computed using the median SpO_{2i} from the 40-second window immediately before the stop time from each motion task and the SaO_2 value from the ABG taken immediately after the same motion task.

Hypoxia Exposure Phase

The metrics were computed using the median SpO_{2i} from a 40-second window, including 35 seconds before and 5 seconds after the i-th reference SaO_2 value (note that SaO_2 readings were taken for the 80%, 83%, 85%, 87%, 90%, 95%, and 100% target values, with the corresponding output of the blood gas analyzer then taken as the reference value). These metrics were also computed for 3 SaO_2 subgroups: severe hypoxia, SaO_2 lower

than 85%; mild hypoxia, SaO_2 from 85% to 89%; and normoxia, SaO_2 equal to or greater than 90%.

Statistical Tests

For both phases, one-way ANOVA followed by the Tukey-Kramer test [24] was used to evaluate differences in the mean bias and the mean absolute bias between groups. The Levene test [25] was used to evaluate differences in the precision between groups. In the movement phase, the distributions between the 15 additional SaO_2 values taken at the middle of the STS motion and those taken at the end were compared via the Wilcoxon test. Significance was considered at $P<.05$.

Sensitivity and Specificity in Detecting Hypoxemia

To evaluate each device's diagnostic accuracy in detecting hypoxemia, we determined the sensitivity, specificity, positive and negative predictive values (PPV and NPV), and accuracy (computed from the error matrix) for identifying values of SaO_2 below 90%. To consider whether device performance would be more reliable if recalibrated, we calculated the area under the receiver operating characteristic (AUROC) curve for each pulse oximeter and computed the same metrics at the optimal operating value. In addition, 95% CIs for all metrics were determined using bootstrapping.

Due to SpO_2 estimation performance issues, Wavelet analysis was removed. Its results can be found in Multimedia Appendix 2.

Results

Participants

Prescreening interviews were performed on 51 volunteers (Consolidated Standards of Reporting Trials [CONSORT] flow diagram in Figure 1). Of these, 1 (2%) volunteer was excluded due to a history of anemia, and 8 (16%) were not able to attend the study session. The remaining 42 (82%) participants attended a study session: 4 (10%) participants presented clinical conditions, evaluated at the start of their session, that were part of the exclusion criteria and would bias the SpO_2 estimations if included (3 [75%] additional anemia cases, evaluated from the first ABG, and 1 [25%] sickle cell trait); for 1 (2%) subject, it was not possible to induce hypoxia; for 2 (5%) subjects, it was not possible to insert an arterial line (in either arm); finally, 2 (5%) subjects had incomplete SpO_2 data for the WristOx2 3150 device. Complete data were therefore obtained from 33 (79%) healthy adults, 18 (55%) women and 15 (45%) men, spread across Fitzpatrick skin types 1-4, with a median age of 29 (SD 24-36) years. The baseline vital sign mean values, measured at the start of each session, were 71 beats per minute (bpm), 15 respirations per minute (rpm), 100% SaO_2, 130/75 mmHg, and a derived BMI of 23.7 kg/m^2 (demographics available in Table 1). Figure 2 shows an exemplar SpO_2 trend for each device for 1 study session. In general, poor estimation performance could be seen during the motion tasks (identified by the brown periods at the top), and all devices (expect Wavelet, whose analysis can be reviewed in Multimedia Appendix 2) followed the SaO_2 desaturation trend (red stars) during the hypoxia phase (from 9:58 AM to 10:14 AM). We discuss their accuracy next.

Figure 1. CONSORT flow diagram of the study. ABG: arterial blood gas; CONSORT: Consolidated Standards of Reporting Trials.

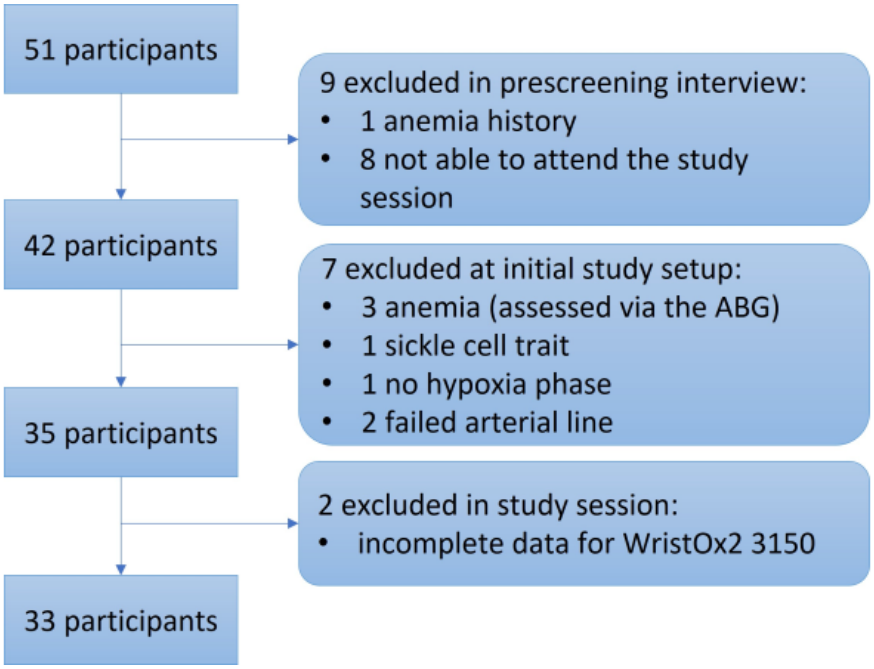


Table 1. Demographics and baseline heart rate, respiration rate, blood pressure, and SaO_2[a] for 33 participants.

Demographics	Mean (median)	Q1[b], Q3[c]
Age (years)	29.0 (31.18)	24.0, 36.0
Sex (female), %	18 (54.5)	N/A[d]
Height (m)	1.70 (1.70)	1.6, 1.8
Weight (kg)	70.0 (70.7)	61.0, 80.0
BMI (kg/m^2)	23.7 (24.3)	21.5, 26.4
Skin tone[e], %		
Type 1	9 (27.3)	N/A
Type 2	15 (45.5)	N/A
Type 3	2 (6.1)	N/A
Type 4	7 (21.2)	N/A
Type 5	0 (0)	N/A
Type 6	0 (0)	N/A
Respiration rate (rpm[f])	15.0 (15.7)	13.0, 18.0
Heart rate (bpm[g])	71.0 (70.9)	62.0, 82.0
SaO_2, %	100.0 (99.6)	100.0, 100.0
Systolic blood pressure (mmHg)	129.5 (133.8)	122.8, 142.8
Diastolic blood pressure (mmHg)	75.0 (77.4)	69.8, 86.3

[a]SaO_2: arterial blood oxygen saturation.

[b]Q_1: first quartile.

[c] Q_3: third quartile.

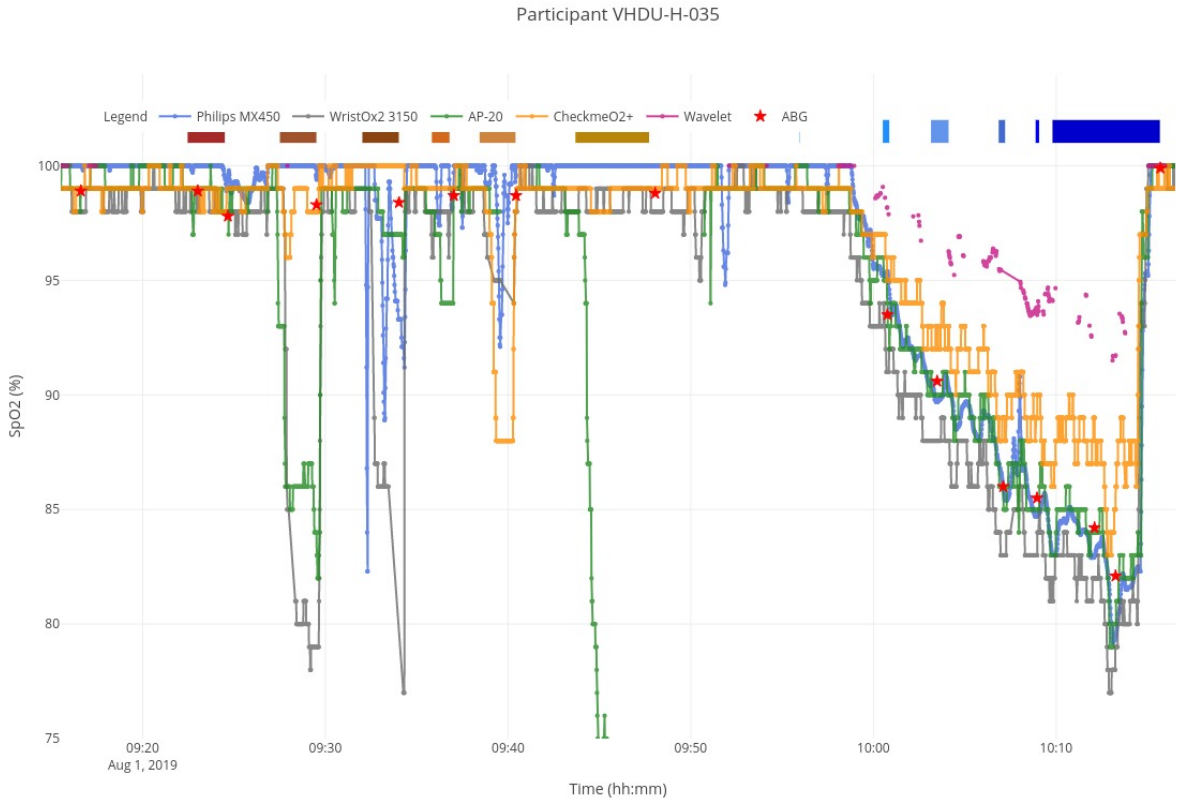
[d]N/A: not applicable.

[e]Fitzpatrick scale.

[f]rpm: respirations per minute.

[g]bpm: beats per minute.

Figure 2. SpO_2 trend for each device during the movement (9:20 AM-9:50 AM) and hypoxia exposure (9:58 AM-10:14 AM) phases of 1 study session. The gold-standard SaO_2, derived from ABG samples, are shown as red stars. The different motion tasks and target desaturation intervals are illustrated by brown and blue rectangles at the top, respectively. Wavelet SpO_2 data are shown for comparison (results can be reviewed in Multimedia Appendix 2). ABG: arterial blood gas; SaO_2: arterial blood oxygen saturation; SpO_2: peripheral oxygen saturation.



SpO_2 Estimation in the Movement Phase

The results of SpO_2 estimation performance metrics for each device are shown in Table 2. The mean bias and precision are further illustrated for each motion task in Figure 3. The number of dropout values was comparable between the finger-worn devices. The RMSE values were below 4% when at rest and for the STS and drinking tasks for all devices. For all other tasks, they were above 4% for at least 1 device.

Table 2. Comparison of accuracy (RMSE[a]) and bias in SpO_2[b] estimation between different motion tasks, for each device, for 33 participants.

Performance metrics	At rest	STS[c]	Rubbing	Tapping	Drinking	Turning pages	Tablet use	*P* value[d]
AP20								
Available SpO_2 points, n	32	30	32	30	31	27	31	N/A[e]
RMSE, % (95% CI)	0.82 (0.55-1.06)	4.68 (1.47-7.72)	11.96 (9.44-14.23)	12.21 (9.31-14.74)	1.96 (1.48-2.46)	8.52 (6.18-10.75)	8.01 (1.15-13.72)	N/A
Mean bias, %	–0.21[f,g]	–0.9[h,i]	–9.91[f,g,j,k]	–9.82[g,i]	–1.45[j]	–6.46	–2.22[k]	<.001
Mean |bias|, %	0.6[k,l,m]	2.15[f,g]	9.91[f,i,j,k]	9.85[g,h,l,n]	1.57[h,j]	6.46[m]	2.56[i,n]	<.001
Precision, %	0.81[f]	4.31[h]	6.91	7.49[f,h,j,k]	1.37[j]	5.57	7.89[k]	<.001
CheckMe O2+								
Available SpO_2 points, n	30	31	31	30	32	32	32	N/A
RMSE, % (95% CI)	1.68 (1.21-2.12)	3.5 (1.49-5.37)	8.45 (5.86-10.88)	3.99 (2.28-5.69)	2.43 (1.9-2.96)	7.83 (5.9-9.8)	4.2 (2.86-5.47)	N/A
Mean bias, %	–1.06	–1.37	–6.19[j]	–2.65[h,j]	–1.93	–6.04[h]	–2.94	.001
Mean |bias|, %	1.29	1.92	6.31	2.71	1.97	6.06	2.98	.005
Precision, %	1.33	3.08	5.84	3.06	1.41	5.11	2.86	.13
Philips MX450								
Available SpO_2 points, n	33	33	33	32	32	32	32	N/A
RMSE, % (95% CI)	1.11 (0.92-1.28)	2.31 (1.9-2.67)	9.49 (7.04-11.86)	7.15 (3.07-10.3)	1.17 (1.0-1.36)	6.64 (3.81-9.03)	1.97 (1.29-2.68)	N/A
Mean bias, %	0.89[f]	1.97[h]	–5.37[f,h,i,j,k]	–1.75[i]	0.84[j]	–3.04	0.4[k]	<.001
Mean |bias|, %	0.97[h]	2.02	6.6[h,j]	3.33	1.06[j]	4.03	1.51	.002
Precision, %	0.63[g,k]	0.78[f]	6.77[f,i,j,k]	7.16[g,h]	0.8[h,j]	6.04	1.82[i]	<.001
WristOx2 3150								
Available SpO_2 points, n	32	33	29	29	32	24	33	N/A
RMSE, % (95% CI)	1.18 (0.84-1.51)	2.33 (1.26-3.41)	9.5 (7.29-11.5)	7.17 (4.66-9.35)	1.27 (0.95-1.57)	6.28 (4.25-8.27)	3.91 (1.49-5.62)	N/A
Mean bias, %	–0.71[f]	–0.4[h]	–7.52[f,h,i,j,k]	–4.56[i]	–0.86[j]	–4.51	–1.81[k]	.002
Mean |bias|, %	0.92[k,l]	1.38[f,g]	7.52[f,i,j,k]	4.69[g,l,m]	1.02[h,j]	4.56	2.02[h,m]	<.001
Precision, %	0.97[f]	2.12[h]	5.98	5.7[f,h,j,k]	0.93[j]	4.35	3.06[k]	.001

[a]RMSE: root-mean-square error.

[b]SpO_2: peripheral oxygen saturation.

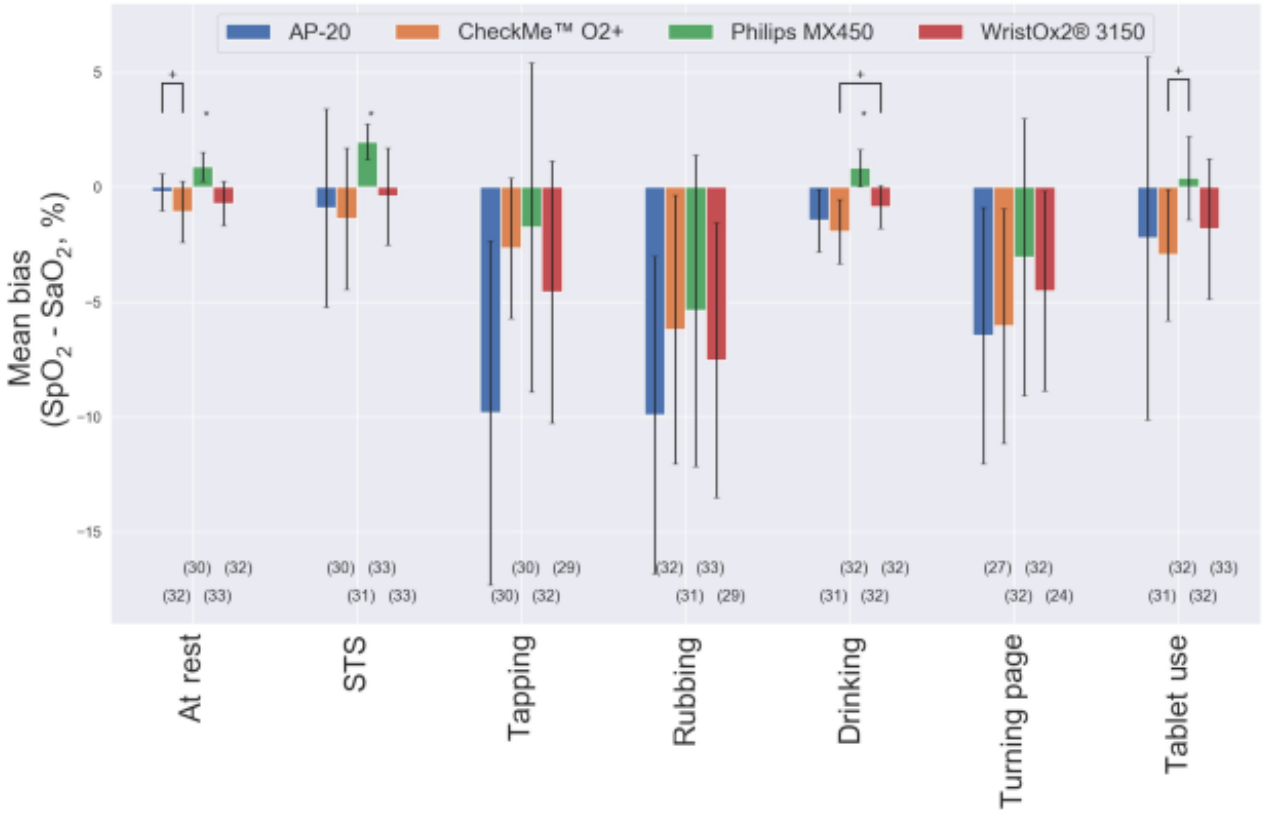
[c]STS: sit-to-stand.

[d]One-way ANOVA followed by the Tukey-Kramer test was used to evaluate differences in the mean bias and mean absolute bias between tasks. The Levene test was used in the case of precision.

[e]N/A: not applicable.

[f-n]Different from each other; for example, for CheckMe O2+, the mean bias of the tapping motion task was different from that of the turning page task and that of the rubbing task (paired differences coded as [j] and [h], respectively).

Figure 3. Comparison of the mean bias (SpO_2–SaO_2) and precision between devices for each movement type. The number of points available per device is presented below each bar. For each task, one-way ANOVA followed by the Tukey test was used to evaluate differences in the mean bias between devices. *Different from other values. +Different from each other. SaO_2: arterial blood oxygen saturation; SpO_2: peripheral oxygen saturation; STS: sit-to-stand.



SpO_2 Estimation in the Hypoxia Exposure Phase

Table 3 compares the SpO_2 estimation performance of the devices across the range of SaO_2 targets of the hypoxia exposure phase, and their Bland-Altman plots (with the mean bias and limits of agreement) can be reviewed in Figure 4. The WristOx2 3150 device underestimated SpO_2 in comparison with SaO_2 by almost 2%. The WristOx2 3150 and CheckMe O2+ devices had the numerically greatest mean absolute bias. However, SaO_2 subgroup analysis (see Figure 5 and Table 4) showed that the WristOx2 3150 device consistently underestimated SpO_2 across the measured range (with an overall mean bias of –1.92% [SD 2.73%]; Table 3), whereas the CheckMe O2+ device overestimated in the severe-hypoxia range and underestimated in the mild-hypoxia and normoxia ranges (Figure 5). However, WristOx2 3150, CheckMe O2+, AP-20, and Philips MX450 showed an overall RMSE below 4% (and below 8% when considering the 95% CI; Table 3), meeting the ISO 80601-2-61:2019 requirement.

Table 3. Comparison of accuracy (RMSE[a]) and mean bias of SpO_2[b] estimation between devices during the hypoxia exposure phase. There were 215 SaO_2 target windows in this phase.

Performance metrics	Philips MX450	CheckMe O2+	WristOx2 3150	AP-20	*P* value
Available SpO_2 points, n	215	207	209	214	N/A[c]
RMSE, % (95% CI)	2.67 (2.31-3.06)	3.20 (2.85-3.56)	3.33 (2.85-3.86)	2.86 (2.44-3.25)	N/A
Mean bias, %	0.49[d]	–0.22	–1.92[e]	–0.3[d]	<.001
Mean |bias|, %	1.92	2.42	2.40	2.00	<.02
Precision, %	2.62[d]	3.16[d]	2.73	2.83	<.02

[a]RMSE: root-mean-square error.

[b]SpO_2: peripheral oxygen saturation.

[c]N/A: not applicable.

[d]Different from each other.

[e]Different from other values.

Figure 4. (a-d) Bland-Altman plots for the Philips MX450, AP-20, CheckMe O2+, and WristOx2 3150 SpO_2 estimation, respectively. The mean bias and limits of agreement values are shown at the left of their respective dashed lines. The solid line represents y=0 (no bias). SaO_2: arterial blood oxygen saturation; SpO_2: peripheral oxygen saturation.

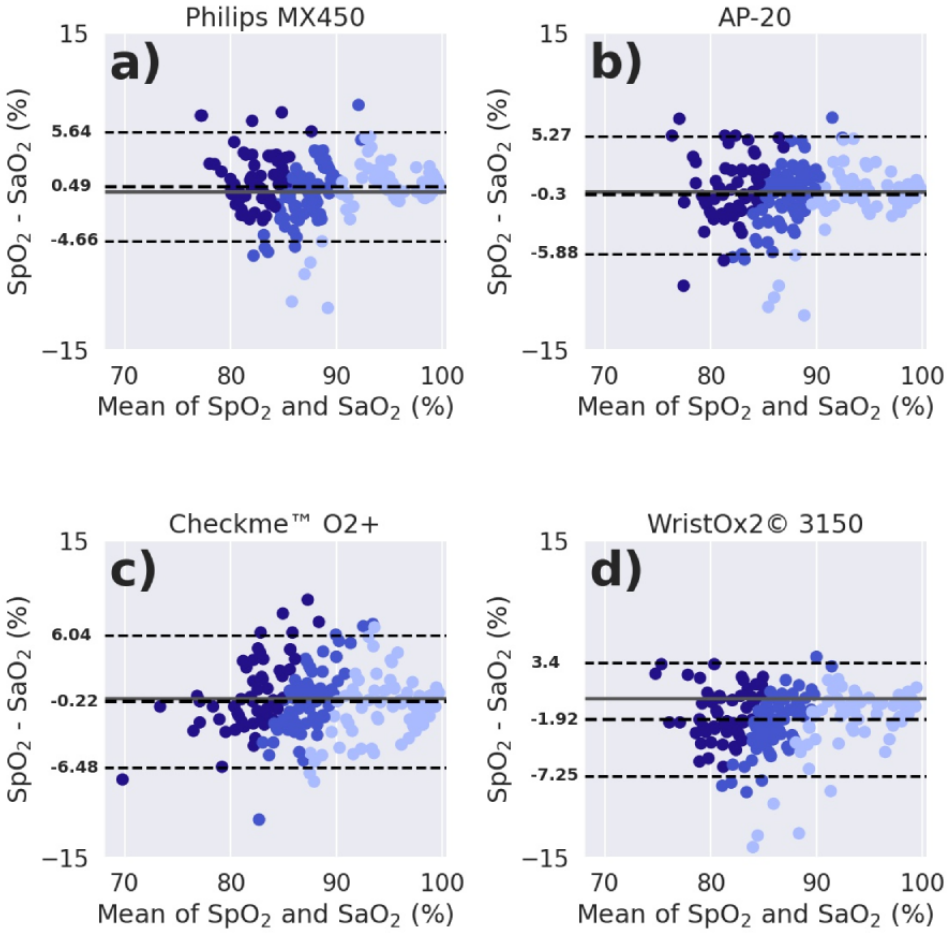


Figure 5. Comparison of the mean bias (SpO_2–SaO_2) and precision between devices for the 3 SaO_2 subgroups: severe hypoxia, SaO_2<85%; mild hypoxia, SaO_2=85%-89%; and normoxia, SaO_2=90%-100%. The number of points available per device is presented below each bar. For each subgroup, one-way ANOVA followed by the Tukey test was used to evaluate differences in the mean bias between devices. *Different from other values. +Different from each other. SaO_2: arterial blood oxygen saturation; SpO_2: peripheral oxygen saturation.

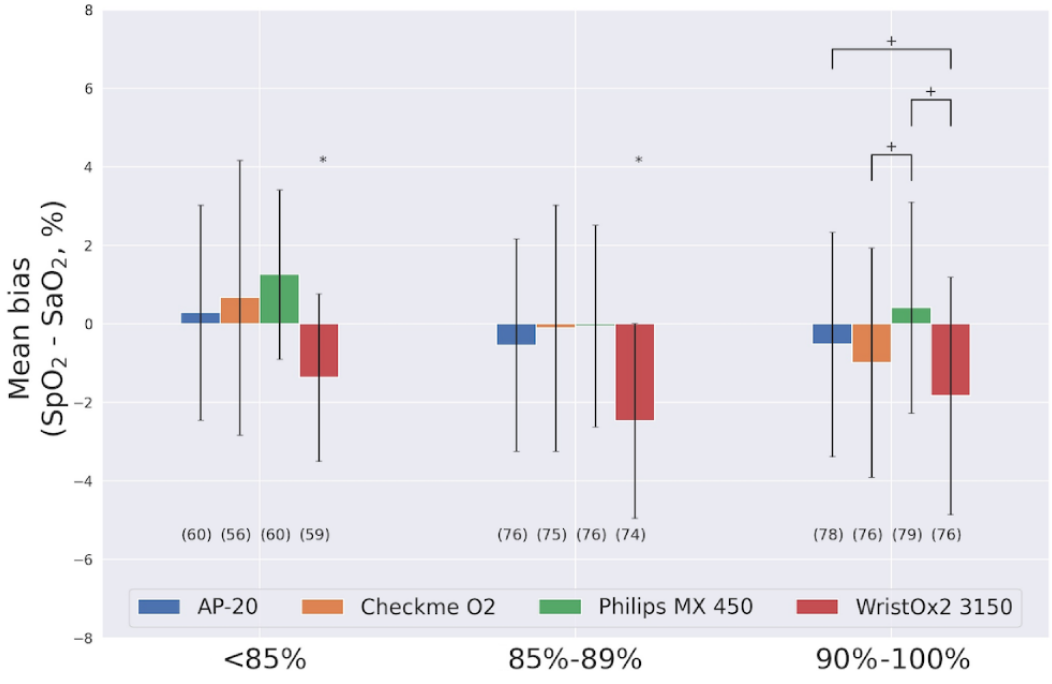


Table 4. Comparison of accuracy (RMSE[a]) and mean bias of the device's SpO_2[b] estimation between 3 SaO_2[c] subgroups: severe hypoxia (SaO_2<85%), mild hypoxia (SaO_2 85%-89%), and normoxia (SaO_2≥90%).

Performance metrics	<85%	85%-89%	90%-100%	*P* value[d]
Total ABGs[e], n	60	76	79	N/A[f]
AP-20				
Available SpO_2 points, n	60	76	78	N/A
RMSE, % (95% CI)	2.99 (2.44-3.57)	2.73 (2.28-3.16)	2.88 (1.88-3.78)	N/A
Mean bias, %	0.28	–0.54	–0.52	.18
Mean |bias|, %	2.33	2.07	1.68	.17
Precision, %	2.74	2.71	2.86	.16
CheckMe O2+				
Available SpO_2 points, n	56	75	76	N/A
RMSE, % (95% CI)	3.52 (2.86-4.18)	3.10 (2.46-3.83)	3.05 (2.54-3.53)	N/A
Mean bias, %	0.67[d]	–0.11[d]	–0.99	.01
Mean |bias|, %	2.74	2.35	2.26	.39
Precision, %	3.5	3.14	2.92	.28
Philips MX 450				
Available SpO_2 points, n	60	76	79	N/A
RMSE, % (95% CI)	2.80 (2.18-3.33)	2.54 (2.08-3.02)	2.70 (1.88-3.56)	N/A
Mean bias, %	1.26	–0.05[g]	0.42[g]	.02
Mean |bias|, %	2.13	1.97	1.72	.44
Precision, %	2.16	2.57	2.69	.24
WristOx2 3150				
Available SpO_2 points, n	59	74	76	N/A
RMSE, % (95% CI)	2.69 (2.28-3.08)	3.49 (2.92-3.99)	3.61 (2.37-4.64)	N/A
Mean bias, %	–1.36	–2.47	–1.83	.06
Mean |bias|, %	2.21	2.83	2.12	.13
Precision, %	2.13	2.48	3.03	.99

[a]RMSE: root-mean-square error.

[b]SpO_2: peripheral oxygen saturation.

[c]SaO_2: arterial blood oxygen saturation.

[d]For each device, one-way ANOVA followed by Tukey test was used to evaluate differences in the mean bias and mean absolute bias between subgroups. The Levene test was used in the case of precision.

[e]ABG: arterial blood gas.

[f]N/A: not applicable.

[g]Different from each other.

Sensitivity and Specificity

Table 5 shows the performance metrics of the pulse oximeters in detecting hypoxemia (SaO_2<90%; AUROC curves available in Figure 6). A total of 128 SaO_2 targets were in the hypoxemia range versus 74 in the normoxia range. At a 90% cut-off, WristOx2 3150 showed significantly better sensitivity (0.97, 95% CI 0.93-0.99) than Philips MX450 (0.86, 95% CI 0.80-0.92). The values for the other metrics were comparable between all devices. All finger-worn devices achieved a good and comparable AUROC curve (≥0.92). Recalibration of the SpO_2 threshold to the optimal operating value resulted in AP-20 achieving significantly higher sensitivity than CheckMe O2+ (0.95 [95% CI 0.91-0.98] vs 0.78 [95% CI 0.71-0.85]). The remaining sensitivity and specificity values were comparable.

Table 5. Performance metrics of each pulse oximeter for detecting hypoxemia (SaO_2[a]<90%). The metrics are shown at a 90% SpO_2[b] cut-off and for the determined optimal SpO_2 cut-off.

Device	Cut-off, %	AUROC[c], mean (95% CI)	Sensitivity, mean (95% CI)	Specificity, mean (95% CI)	PPV[d], mean (95% CI)	NPV[e], mean (95% CI)	Accuracy[f], mean (95% CI)
90% SpO_2 (%) cut-off							
Philips MX 450	90.0	N/A[g]	0.86 (0.80-0.92)	0.93 (0.87-0.99)	0.96 (0.92-0.99)	0.79 (0.71-0.88)	0.89 (0.84-0.93)
CheckMe O2+	90.0	N/A	0.87 (0.81-0.93)	0.85 (0.76-0.93)	0.91 (0.86-0.96)	0.80 (0.71-0.88)	0.87 (0.82-0.91)
WristOx2 3150	90.0	N/A	0.97 (0.93-0.99)	0.80 (0.70-0.89)	0.89 (0.84-0.94)	0.94 (0.87-0.99)	0.91 (0.86-0.95)
AP-20	90.0	N/A	0.91 (0.85-0.95)	0.89 (0.82-0.96)	0.94 (0.89-0.98)	0.85 (0.76-0.92)	0.90 (0.86-0.94)
Optimal SpO_2 (%) cut-off obtained via AUROC analysis[h]							
Philips MX 450	90.7	0.94 (0.90-0.98)	0.97 (0.94-0.99)	0.86 (0.78-0.94)	0.93 (0.88-0.97)	0.94 (0.88-0.99)	0.93 (0.90-0.97)
CheckMe O2+	89.0	0.92 (0.87-96)	0.78 (0.71-0.85)	0.88 (0.80-0.95)	0.92 (0.86-0.97)	0.70 (0.60-0.79)	0.82 (0.76-0.87)
WristOx2 3150	88.0	0.94 (0.89-97)	0.88 (0.82-0.94)	0.86 (0.78-0.94)	0.92 (0.87-0.96)	0.81 (0.72-0.89)	0.88 (0.83-0.92)
AP-20	91.0	0.94 (0.89-98)	0.95 (0.91-0.98)	0.84 (0.75-0.92)	0.91 (0.86-0.96)	0.91 (0.84-0.97)	0.91 (0.87-0.95)

[a]SaO_2: arterial blood oxygen saturation.

[b]SpO_2: peripheral oxygen saturation.

[c]AUROC: area under the receiver operating characteristic.

[d]PPV: positive predictive value.

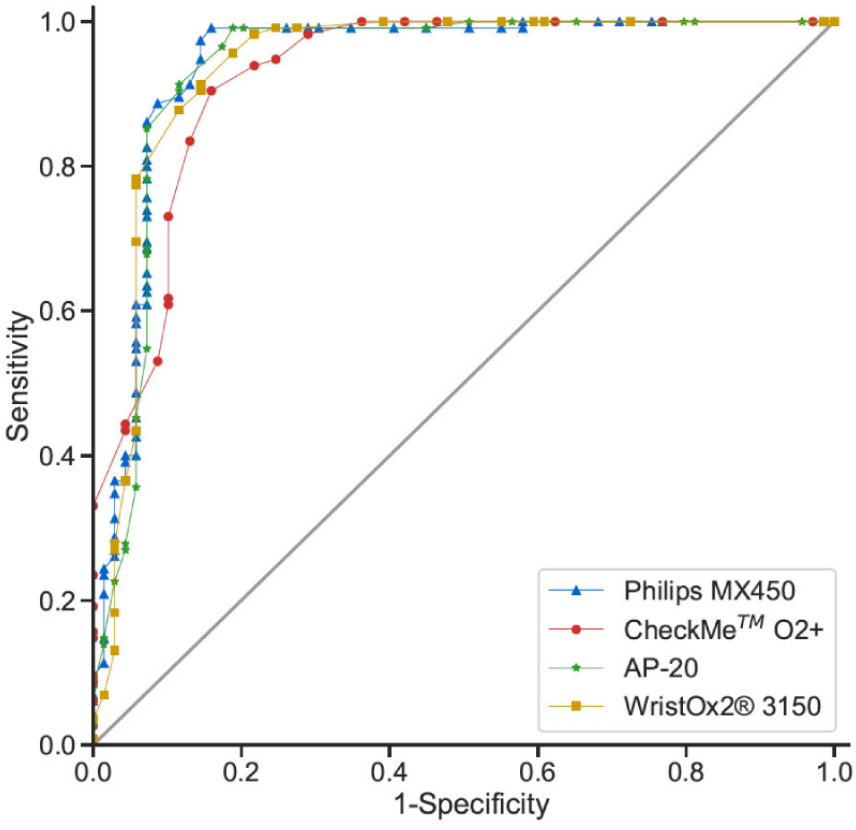
[e]NPV: negative predictive value.

[f]Accuracy = (True positives + True negatives)/n, where n is the total number of examples.

[g]N/A: not applicable.

[h]The optimal SpO_2 cut-off is the best compromise between sensitivity and specificity to detect hypoxemia (SaO_2<90%).

Figure 6. ROC curves in detecting hypoxemia (SaO_2<90%) during the hypoxia exposure phase. ROC: area under the receiver operating characteristic; SaO_2: arterial blood oxygen saturation.



Discussion

Principal Findings

Several studies have been published on both the usefulness and the potential issues of pulse oximetry in the clinical setting using nonambulatory devices. In this study, we compared the performance of wearable pulse oximeters and 1 nonambulatory pulse oximeter using gold-standard arterial blood samples drawn from healthy adult participants. Availability of waveform data was a requirement that limited the selection of devices. Our provision of waveform data will allow clinical staff to assess the reliability of the signal. However, a risk with all continuous-monitoring systems is that they increase the burden on clinical teams by providing excess data. Further work is required to determine the usefulness of these systems in clinical practice and how continuous-monitoring data should be summarized in the electronic patient record.

In tests of finger-based devices, WristOx2 3150 significantly underestimated SaO_2 (mean bias –1.92% [SD 2.73%]; Table 3) when compared with the other wearables. Nevertheless, all finger-based probes showed a similar mean absolute bias (about 2%) and RMSE (about 3%). Overall, all finger-worn wearable pulse oximeters achieved good sensitivity (≥0.87) and specificity (≥0.80), comparable to the standard nonambulatory device, in detecting hypoxemia (Table 5). Given that WristOx2 3150 underestimates SaO_2, it presented higher sensitivity (0.97, 95% CI 0.93-0.99) at the cost of a lower specificity value (0.80, 95% CI 0.70-0.89). This underestimation explains why recalibration by 2% achieves the optimal operating point. The remaining devices only required a change in the threshold by 1% at their optimal operating point.

From the 7 motion tasks, tapping, rubbing, turning book pages, and using a tablet challenged the finger-based wearable devices the most (the first 2 are also analyzed by Louie et al [11] and Barker and Shah [26]), resulting in an RMSE above 4% in at least 1 device. The mean bias at rest, STS, and drinking motions was comparable (<4%; Table 2).

Limitations

The sample size calculation for our study was based on the ISO 80601-2-61:2019 guidelines to evaluate the accuracy of pulse oximeters in detecting changes in SpO_2, not to identify differences in performance between pulse oximeters and between activities. The study was not designed to generalize results to the wider population, for example, for patients with darker skin types or with acute illness.

We chose to sample ABGs at the end of each task to avoid accidental removal of the cannula. However, it became clear during our study that the ABGs could have been sampled while the motion task was occurring, perhaps better representing that interval reference SaO_2. Preliminary analysis of the difference between ABGs taken immediately after the STS motion task and those taken at the midpoint of that motion, for 15 patients, showed that the SaO_2 dropped by an average of 1.87% (SD 0.87%, $P<.001$ between the 2 SaO_2 sample sets), indicating that the SpO_2 value might change between the time used to compute the SpO_2 estimates and that of the ABG samples taken after the exercise. Our hypothesis was that the STS task would be the motion task with the greatest effect on the participants' SaO_2. However, the error in the SpO_2 estimates from the wearable devices during motion was much larger, so this correction would not have changed our findings.

Conclusion

The accuracy of SpO_2 estimation by finger-worn pulse oximeters was within that required by the ISO 80601-2-61:2019 guideline (≤4%). The accuracy was degraded by motion but not more than that with usual-care bedside monitors. All finger-worn pulse oximeters were capable of detecting hypoxemia, their performance being comparable to that used in nonambulatory standard care.

Our findings support the use of wearable, finger-based, wireless transmission–mode pulse oximeters to detect the onset of clinical deterioration in the hospital, possibly earlier than intermittent vital-sign measurements. The continuous assessment of SpO_2, especially values below 90%, may be helpful to manage the care of ambulatory in-hospital patients who have been infected with the SARS-CoV-2 virus [27]. Further work is required to assess the impact of AMSs on patient outcomes, both during the COVID-19 pandemic and beyond.

Acknowledgments

This work was supported by the National Institute for Health Research (NIHR) Biomedical Research Centre (BRC), Oxford, UK. MS, LY, SV, PW, and CA were funded/supported by the NIHR BRC. The views expressed are those of the authors and not necessarily those of the National Health Service (NHS), the NIHR, or the Department of Health. MAFP was funded by a Drayson research fellowship. AS is currently supported by an NIHR doctoral research fellowship (NIHR-DRF-2017-10-094).

Conflicts of Interest

PW and LT report significant grants from the National Institute of Health Research (NIHR), UK, and the NIHR Biomedical Research Centre (BRC), Oxford, UK, during the conduct of the study. PW and LT report modest grants and personal fees from Sensyne Health, outside the submitted work. LT works part-time for Sensyne Health and holds share options in the company. PW also holds shares in the company.

Multimedia Appendix 1
Analysis of the pulse rate estimation in wearable pulse oximeters.
[PDF File (Adobe PDF File), 1140 KB - jmir_v24i2e28890_app1.pdf]

Multimedia Appendix 2
Results of Wavelet's SpO_2 estimation.
[PDF File (Adobe PDF File), 540 KB - jmir_v24i2e28890_app2.pdf]

References

1. Sun L, Joshi M, Khan SN, Ashrafian H, Darzi A. Clinical impact of multi-parameter continuous non-invasive monitoring in hospital wards: a systematic review and meta-analysis. J R Soc Med 2020 Jun;113(6):217-224 [FREE Full text] [doi: 10.1177/0141076820925436] [Medline: 32521195]
2. McQuillan P, Pilkington S, Allan A, Taylor B, Short A, Morgan G, et al. Confidential inquiry into quality of care before admission to intensive care. BMJ 1998 Jun 20;316(7148):1853-1858 [FREE Full text] [doi: 10.1136/bmj.316.7148.1853] [Medline: 9632403]
3. Leenen JPL, Leerentveld C, van Dijk JD, van Westreenen HL, Schoonhoven L, Patijn GA. Current evidence for continuous vital signs monitoring by wearable wireless devices in hospitalized adults: systematic review. J Med Internet Res 2020 Jun 17;22(6):e18636 [FREE Full text] [doi: 10.2196/18636] [Medline: 32469323]
4. Royal College of Physicians. National Early Warning Score (NEWS) 2. Standardising the Assessment of Acute-Illness Severity in the NHS. London: Royal College of Physicians; Dec 19, 2017.
5. Downey CL, Tahir W, Randell R, Brown JM, Jayne DG. Strengths and limitations of early warning scores: a systematic review and narrative synthesis. Int J Nurs Stud 2017 Nov;76:106-119. [doi: 10.1016/j.ijnurstu.2017.09.003] [Medline: 28950188]
6. Downey C, Randell R, Brown J, Jayne DG. Continuous versus intermittent vital signs monitoring using a wearable, wireless patch in patients admitted to surgical wards: pilot cluster randomized controlled trial. J Med Internet Res 2018 Dec 11;20(12):e10802 [FREE Full text] [doi: 10.2196/10802] [Medline: 30538086]
7. Joshi M, Ashrafian H, Aufegger L, Khan S, Arora S, Cooke G, et al. Wearable sensors to improve detection of patient deterioration. Expert Rev Med Devices 2019 Feb;16(2):145-154. [doi: 10.1080/17434440.2019.1563480] [Medline: 30580650]
8. Watkins T, Whisman L, Booker P. Nursing assessment of continuous vital sign surveillance to improve patient safety on the medical/surgical unit. J Clin Nurs 2016 Jan;25(1-2):278-281 [FREE Full text] [doi: 10.1111/jocn.13102] [Medline: 26531215]
9. Weller RS, Foard KL, Harwood TN. Evaluation of a wireless, portable, wearable multi-parameter vital signs monitor in hospitalized neurological and neurosurgical patients. J Clin Monit Comput 2018 Oct;32(5):945-951. [doi: 10.1007/s10877-017-0085-0] [Medline: 29214598]
10. Paul JE, Chong MA, Buckley N, Harsha P, Shanthanna H, Tidy A, et al. Vital sign monitoring with continuous pulse oximetry and wireless clinical notification after surgery (the VIGILANCE pilot study): a randomized controlled pilot trial. Pilot Feasibility Stud 2019 Feb 26;5(1):36 [FREE Full text] [doi: 10.1186/s40814-019-0415-8] [Medline: 30858986]
11. Louie A, Feiner J, Bickler P, Rhodes L, Bernstein M, Lucero J. Four types of pulse oximeters accurately detect hypoxia during low perfusion and motion. Anesthesiology 2018 Mar;128(3):520-530 [FREE Full text] [doi: 10.1097/ALN.0000000000002002] [Medline: 29200008]
12. Appelboom G, Camacho E, Abraham ME, Bruce SS, Dumont EL, Zacharia BE, et al. Smart wearable body sensors for patient self-assessment and monitoring. Arch Public Health 2014 Aug;72(1):28 [FREE Full text] [doi: 10.1186/2049-3258-72-28] [Medline: 25232478]
13. Petterson MT, Begnoche VL, Graybeal JM. The effect of motion on pulse oximetry and its clinical significance. Anesth Analg 2007 Dec;105(6 Suppl):S78-S84. [doi: 10.1213/01.ane.0000278134.47777.a5] [Medline: 18048903]
14. Weenk M, van Goor H, Frietman B, Engelen LJ, van Laarhoven CJ, Smit J, et al. Continuous monitoring of vital signs using wearable devices on the general ward: pilot study. JMIR Mhealth Uhealth 2017 Jul 05;5(7):e91 [FREE Full text] [doi: 10.2196/mhealth.7208] [Medline: 28679490]
15. Verrillo SC, Cvach M, Hudson KW, Winters BD. Using continuous vital sign monitoring to detect early deterioration in adult postoperative inpatients. J Nurs Care Qual 2019;34(2):107-113. [doi: 10.1097/NCQ.0000000000000350] [Medline: 30095509]
16. Seshadri DR, Davies EV, Harlow ER, Hsu JJ, Knighton SC, Walker TA, et al. Wearable sensors for COVID-19: a call to action to harness our digital infrastructure for remote patient monitoring and virtual assessments. Front Digit Health 2020 Jun 23;2:8 [FREE Full text] [doi: 10.3389/fdgth.2020.00008] [Medline: 34713021]
17. O'Carroll O, MacCann R, O'Reilly A, Dunican EM, Feeney ER, Ryan S, et al. Remote monitoring of oxygen saturation in individuals with COVID-19 pneumonia. Eur Respir J 2020 Aug;56(2):2001492 [FREE Full text] [doi: 10.1183/13993003.01492-2020] [Medline: 32616588]

18. Greenhalgh T, Knight M, A'Court C, Buxton M, Husain L. Management of post-acute covid-19 in primary care. BMJ 2020 Aug 11;370:m3026. [doi: 10.1136/bmj.m3026] [Medline: 32784198]
19. Cohen JF, Korevaar DA, Gatsonis CA, Glasziou PP, Hooft L, Moher D, STARD Group. STARD for Abstracts: essential items for reporting diagnostic accuracy studies in journal or conference abstracts. BMJ 2017 Aug 17;358:j3751. [doi: 10.1136/bmj.j3751] [Medline: 28819063]
20. Areia C, Vollam S, Piper P, King E, Ede J, Young L, et al. Protocol for a prospective, controlled, cross-sectional, diagnostic accuracy study to evaluate the specificity and sensitivity of ambulatory monitoring systems in the prompt detection of hypoxia and during movement. BMJ Open 2020 Jan 12;10(1):e034404 [FREE Full text] [doi: 10.1136/bmjopen-2019-034404] [Medline: 31932393]
21. Haahr M. What's This Fuss about True Randomness?. URL: https://www.random.org [accessed 2021-02-28]
22. Rowland MJ, Ezra M, Winkler A, Garry P, Lamb C, Kelly M, et al. Calcium channel blockade with nimodipine reverses MRI evidence of cerebral oedema following acute hypoxia. J Cereb Blood Flow Metab 2019 Feb;39(2):285-301 [FREE Full text] [doi: 10.1177/0271678X17726624] [Medline: 28857714]
23. Fitzpatrick TB. The validity and practicality of sun-reactive skin types I through VI. Arch Dermatol 1988 Jun;124(6):869-871. [doi: 10.1001/archderm.124.6.869] [Medline: 3377516]
24. Haynes W. Tukey's test. In: Dubitzky W, Wolkenhauer O, Cho KH, Yokota H, editors. Encyclopedia of Systems Biology. New York, NY: Springer; 2013:2303-2304.
25. NIST/SEMATECH. NIST/SEMATECH e-Handbook of Statistical Methods. 2012. URL: https://www.itl.nist.gov/div898/handbook/ [accessed 2020-02-03]
26. Barker SJ, Shah NK. Effects of motion on the performance of pulse oximeters in volunteers. Anesthesiology 1996 Oct;85(4):774-781 [FREE Full text] [doi: 10.1097/00000542-199610000-00012] [Medline: 8873547]
27. Oxford University Hospitals. New Wearable Technology to Monitor COVID-19 Patients. URL: https://www.ouh.nhs.uk/covid-19/news/article.aspx?id=1215 [accessed 2022-02-03]

Abbreviations

ABG: arterial blood gas
AMS: ambulatory monitoring system
AUROC: area under the receiver operating characteristic
CONSORT: Consolidated Standards of Reporting Trials
BLE: Bluetooth Low Energy
bpm: beats per minute
ECG: electrocardiogram
FiO_2: fraction of inspired oxygen
ISO: International Organization for Standardization
NPV: negative predictive value
PPV: positive predictive value
RMSE: root-mean-square error
ROC: receiver operating characteristic
rpm: respirations per minute
SaO_2: arterial blood oxygen saturation
SpO_2: peripheral oxygen saturation
STS: sit-to-stand

Edited by R Kukafka; submitted 21.03.21; peer-reviewed by C Downey, M Knight; comments to author 21.04.21; revised version received 04.05.21; accepted 21.11.21; published 15.02.22.

Please cite as:
Santos M, Vollam S, Pimentel MAF, Areia C, Young L, Roman C, Ede J, Piper P, King E, Harford M, Shah A, Gustafson O, Tarassenko L, Watkinson P
The Use of Wearable Pulse Oximeters in the Prompt Detection of Hypoxemia and During Movement: Diagnostic Accuracy Study
J Med Internet Res 2022;24(2):e28890
URL: https://www.jmir.org/2022/2/e28890
doi:10.2196/28890
PMID:35166690

Original Paper

Behavioral Efficacy of a Sexual Health Mobile App for Men Who Have Sex With Men: Randomized Controlled Trial of Mobile Messaging for Men

Patrick Sean Sullivan[1], DVM, PhD; Rob Stephenson[2], PhD; Sabina Hirshfield[3], PhD; Cyra Christina Mehta[4,5], MSPH, PhD; Ryan Zahn[1], MSPH, MBA; Jose A Bauermeister[6], MPH, PhD; Keith Horvath[7], PhD; Mary Ann Chiasson[8,9], DrPH; Deborah Gelaude[10], MA; Shelby Mullin[1], MPH; Martin J Downing Jr[11], PhD; Evelyn Jolene Olansky[10,12], MPH; Sarah Wiatrek[1], MPH; Erin Q Rogers[1,2], MPH; Eli Rosenberg[13], PhD; Aaron J Siegler[1], MPH, PhD; Gordon Mansergh[10], PhD

[1]Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, United States

[2]Department of Systems, Population, and Leadership, School of Nursing, University of Michigan, Ann Arbor, MI, United States

[3]Special Treatment and Research Program, Department of Medicine, The State University of New York Downstate Health Sciences University, Brooklyn, NY, United States

[4]Division of Infectious Diseases, Department of Medicine, Emory University School of Medicine, Atlanta, GA, United States

[5]Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, NY, United States

[6]Department of Family and Community Health, School of Nursing, University of Pennslyvania, Philadelphia, PA, United States

[7]Department of Psychology, San Diego State University, San Diego, CA, United States

[8]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, United States

[9]Division of Infectious Diseases, Department of Medicine, Columbia University Irving Medical Center, New York, NY, United States

[10]HIV Research Branch, Division of HIV Prevention, Centers for Disease Control and Prevention, Atlanta, GA, United States

[11]Department of Psychology, Lehman College, City University of New York, Bronx, NY, United States

[12]Social & Scientific Systems, Inc, DLH Holdings Company, Atlanta, GA, United States

[13]Department of Epidemiology, School of Public Health, University at Albany, State University of New York, Albany, NY, United States

Corresponding Author:
Patrick Sean Sullivan, DVM, PhD
Department of Epidemiology
Rollins School of Public Health
Emory University
1518 Clifton Road NE
Atlanta, GA, 30322
United States
Phone: 1 404 210 6039
Email: pssulli@emory.edu

Abstract

Background: Gay, bisexual, and other men who have sex with men (GBMSM) face the highest burden of HIV in the United States, and there is a paucity of efficacious mobile health (mHealth) HIV prevention and care interventions tailored specifically for GBMSM. We tested a mobile app combining prevention messages and access to core prevention services for GBMSM.

Objective: This study aims to measure the efficacy of the Mobile Messaging for Men (M-cubed) app and related services to increase HIV prevention and care behaviors in diverse US GBMSM.

Methods: We conducted a randomized open-label study with a waitlist control group among GBMSM in 3 groups (low-risk HIV-negative group, high-risk HIV-negative group, and living-with-HIV [LWH] group) recruited online and in venues in Atlanta, Detroit, and New York City. Participants were randomly assigned to receive access to the app immediately or at 9 months after randomization. The app provided prevention messages in 6 domains of sexual health and offered ordering of at-home HIV and sexually transmitted infection test kits, receiving preexposure prophylaxis (PrEP) evaluations and navigation, and service locators. Serostatus- and risk-specific prevention outcomes were evaluated at baseline, at the end of the intervention period, and at 3, 6, and 9 months after the intervention period.

Results: In total, 1226 GBMSM were enrolled and randomized; of these 611 (49.84%) were assigned to the intervention group and 608 (99.51%) were analyzed, while 615 (50.16%) were assigned to the control group and 612 (99.51%) were analyzed. For high-risk GBMSM, allocation to the intervention arm was associated with higher odds of HIV testing during the intervention period (adjusted odds ratio [aOR] 2.02, 95% CI 1.11-3.66) and with higher odds of using PrEP in the 3 months after the intervention period (aOR 2.41, 95% CI 1.00-5.76, *P*<.05). No changes in HIV prevention or care were associated with allocation to the intervention arm for the low-risk HIV-negative and LWH groups.

Conclusions: Access to the M-cubed app was associated with increased HIV testing and PrEP use among high-risk HIV-negative GBMSM in 3 US cities. The app could be made available through funded HIV prevention providers; additional efforts are needed to understand optimal strategies to implement the app outside of the research setting.

Trial Registration: ClinicalTrials.gov NCT03666247; https://clinicaltrials.gov/ct2/show/NCT03666247

International Registered Report Identifier (IRRID): RR2-10.2196/16439

(J Med Internet Res 2022;24(2):e34574) doi:10.2196/34574

KEYWORDS

HIV prevention; mHealth; tool; video; randomized clinical trial; app; prevention; HIV; PrEP; STI; testing; behavior; efficacy; men who have sex with men; MSM; sexuality; gay; bisexual; United States

Introduction

The ambitious goal of the US Ending the HIV Epidemic Initiative to reduce new HIV diagnoses by 90% by 2030 [1] will not be reached without reducing new HIV infections in gay, bisexual, and other men who have sex with men (GBMSM). GBMSM comprised 54% of annual new HIV diagnoses in 2019 in the United States [2] and face the highest burden of HIV in the United States [3]. However, there are few efficacious or promising mobile health (mHealth) HIV prevention and care interventions tailored for subgroups of GBMSM [4,5]. New HIV prevention tools should address the needs of GBMSM, young GBMSM aged 15-24 years (YGBMSM) [6], and YGBMSM of color, for whom HIV incidence is the highest [7-9]. Currently, HIV prevention services are underused by GBMSM: in 2018, just over half (56%) reported being tested in the past 12 months, high proportions (76% of HIV-positive and 66% of HIV-negative GBMSM) reported recent condomless receptive anal intercourse, and few (<20%) reported preexposure prophylaxis (PrEP) use [10,11]. Statistical models demonstrate that high levels of coverage of prevention services will be required to substantially reduce HIV incidence among GBMSM [12,13] and increased use of routine prevention services (eg, HIV testing) might increase uptake of biomedical interventions, such as PrEP [14]. A package of HIV prevention interventions is needed to maximally reduce the risks of infection, and uptake of multiple prevention services by a substantial proportion of the GBMSM population is needed to reduce HIV incidence among GBMSM.

A growing body of research has suggested that mobile phone apps are a promising environment to offer tailored and on-demand prevention services for GBMSM [5,6,15-19]. GBMSM are open to receiving prevention information and resources via mobile apps [20]. Communicating prevention messages through mobile apps (mobile messaging) might enhance intervention uptake, because it allows messaging to a wide audience of GBMSM, including rural GBMSM, who use sexual health services and PrEP at lower rates than GBMSM in urban areas [10,11,21]. Younger GBMSM might be especially interested in using mobile technology to receive health information [22].

HIV prevention and care interventions for GBMSM have most often been provided through in-person sessions with behavior change interventions, are often focused on specific subgroups of GBMSM (eg, high-risk HIV-negative episodic substance-using GBMSM, Black GBMSM with HIV-negative or unknown HIV serostatus, HIV-positive clinic patients) [23] and often target only 1 element of comprehensive prevention—for example, condom use, medication adherence, or PrEP uptake [4]. In a serostatus-neutral framework [24], we must evaluate interventions that address multiple prevention behaviors (eg, HIV testing, sexually transmitted infection [STI] testing, condom use, PrEP uptake, and medication adherence), many of which are relevant to GBMSM with HIV and those at risk for HIV infection.

To address these needs, we adapted an existing HIV prevention app designed for GBMSM, HealthMindr, to add and evaluate tailored prevention messaging [25-27]. The methods for the app content development specific to this study have been described elsewhere [28]. We aimed to develop a mobile app that can address multiple prevention and care needs for GBMSM, both those with HIV and those at risk for HIV acquisition (ie, HIV negative). We developed messaging for HIV-negative GBMSM that would be relevant for both high- and low-risk MSM. Messages were drawn from existing mHealth interventions [29,30] and adapted through a theory-driven approach with stakeholder input, as previously described [28]. Messages were provided through in-app content and videos. We used the core HealthMindr app [25,26] to offer a suite of prevention services, including self-screening for HIV and STI risk; a scheduling and reminder system for routine HIV and STI testing; a PrEP eligibility screener; a nonoccupational postexposure prophylaxis (nPEP) risk assessment tool; an ordering platform for delivery of at-home HIV- and STI-screening kits and of condoms and lubricants; and service locators for HIV and STI testing, PrEP, nPEP, and HIV treatment and care. The app included branding associating the app with Emory University. Content was frozen during the trial. Participants accessed the app from Apple App

Store or Google Play Store and received a study-specific activation code.

We aimed to evaluate the use and efficacy of the combined prevention messaging and app service components as a public health intervention strategy to reduce risks of HIV acquisition and improve HIV care outcomes among GBMSM. We hypothesized that exposure to the message delivery platform within a comprehensive HIV prevention app would result in improvements in participants' self-reported sexual health, prevention, and care behaviors, compared to GBMSM without access to the app.

Methods

Study Design

Data were collected from January 24, 2018, to October 31, 2019. The methods for the study have been previously described [28] and are summarized here. Eligibility criteria are summarized in Table 1. Men were classified as (1) HIV seropositive, (2) HIV seronegative at high risk (any anal sex not protected by condoms or not taking PrEP, as prescribed, in the past 3 months), and (3) HIV seronegative at low risk (anal sex protected by condoms in the past 3 months or taking PrEP, as prescribed, in the past 3 months, ie, adherent PrEP users). Owning a cell phone was also required, and smartphone literacy was thus a de facto eligibility criterion. This research was reviewed and approved by the Emory University institutional review board (protocol IRB00087684) on June 13, 2016. We collected individual informed consent in person at the study enrollment visit. Consenting participants were randomly assigned to either the immediate intervention group or the waitlist control group in a 1:1 ratio within 36 strata based on the 3 serostatus/risk groups, 3 cities, race/ethnicity (non-Hispanic White vs other race/ethnicity), and phone type (Android vs iOS) [31]. Participants in the intervention group received one 2-sentence message every other day and a 1-minute video weekly, in addition to access to HIV and STI kit ordering and informational resources.

At in-person baseline visits, participants assigned to the intervention group downloaded and received orientation to the study app and received access to the mobile app and prevention messages based on their risk profiles. Participant contact and retention activities were supported through the Emory Study Management and Retention Tool (SMART). The intervention was provided over 3 months, after which the app was deactivated. Waitlist control participants were not provided access to the intervention app at the baseline visit but continued to receive quarterly surveys. An overview of the study timeline and complete survey instruments have been published and are available online [28].

There were multiple prevention outcomes, which varied among the 3 risk groups, as previously reported [28]. The specific outcomes and measures used are summarized in Table 2.

Table 1. Eligibility criteria for a randomized controlled trial of the Mobile Messaging for Men intervention in Atlanta, Detroit, and New York City (2018-2019).

Criterion	Eligible level
Age (years)	• ≥18
Sex at birth	• Assigned as male
Current gender identity	• Male
Sexual risk behaviors	• Reports anal intercourse with a male partner in the past 12 months
City of residence	• Resident of Atlanta, Detroit, or New York City
Other eligibility criteria	• Plans to stay in the city area for the next 9 months • Owns and uses an Android or iOS smartphone • Is able to read and understand English without assistance

Table 2. Outcomes and measures used in the evaluation of Mobile Messaging for Men, a mobile health intervention to promote HIV prevention and care behaviors among GBMSM[a] in Atlanta, Detroit, and New York City (2018-2019).

Behavior	Group assessed	Definition
STI[b] testing	• All GBMSM	STI test[c] in the 3 months before the survey
HIV testing	• Low-risk HIV-negative GBMSM • High-risk HIV-negative GBMSM	Self-reported[d] HIV test in the 3 months before the survey
Current PrEP[e] use	• Low-risk HIV-negative GBMSM • High-risk HIV-negative GBMSM	Self-reported[d] being on PrEP as of the day of the survey
PrEP adherence	• All low- and high-risk HIV-negative GBMSM reporting PrEP use	Self-reported[d] taking at least 25/30 daily pills in the prior 30 days
Current ART[f] use	• All LWH[g] GBMSM	Self-reported[d] being on ART as of the day of the survey
ART adherence	• All LWH GBMSM on ART	Self-reported[d] taking ART, as prescribed, on ≥25 of the past 30 days
Anal sex not protected by PrEP or condoms	• Low-risk HIV-negative GBMSM • High-risk HIV-negative GBMSM	Self-reported anal sex with a main or casual partner when the participant was not on PrEP AND the insertive partner did not use a condom from start to finish[d]
Anal sex not protected by PrEP or condoms	• LWH GBMSM	Self-reported anal sex with a main or casual partner when the insertive partner did not use a condom from start to finish[d]

[a]GBMSM: gay, bisexual, and other men who have sex with men.

[b]STI: sexually transmitted infection.

[c]Includes Mobile Messaging for Men care kit orders with results.

[d]Self-report measures were previously validated as part of the American Men's Internet Survey [32].

[e]PrEP: preexposure prophylaxis.

[f]ART: antiretroviral therapy.

[g]LWH: living with HIV.

Data Analysis

Analyses were stratified by the baseline HIV status/risk group: high-risk HIV-negative, low-risk HIV-negative, and living with HIV (LWH) groups. Counts and relative frequencies were used to describe sociodemographics and sexual history at baseline by intervention status (intervention group vs control group) and HIV status/risk group, and chi-square tests assessed for differences in these characteristics by intervention status (Table 3). To account for within-person repeated measures and randomization strata, separate mixed-effect logistic models assessed for association over time of the intervention with each outcome.

A post hoc analysis was conducted to compare control participants with intervention participants who had at least 30 minutes of intervention use in order to assess the potentially mediating effect of the time spent on the app; 30 minutes was the estimated time required to read and rate all written messages and view the video messages. A second post hoc analysis was conducted because not all participants who ordered and received at-home STI test kits returned them. In this second analysis, we defined the intention to test for STIs as either ordering an STI test kit through the mobile app or reporting having taken an STI test. Result reporting was prepared in accordance with Consolidated Standards of Reporting Trials (CONSORT) guidelines [33].

Table 3. Sociodemographic and baseline behavioral characteristics of MSM[a] in the Mobile Messaging for Men study sample (2018-2019).

Characteristic	Overall (N=1220)	High-risk[b] HIV-negative (n=427)		Low-risk[c] HIV-negative (n=410)		HIV-positive (n=383)	
	n (%)	n (%)	*P* value[d] (intervention vs control)	n (%)	*P* value (intervention vs control)	n (%)	*P* value (intervention vs control)
Race/ethnicity[e]							
MSM of color	709 (58.11)	205 (48.0)	.89	204 (49.8)	.92	300 (78.3)	.84
White	510 (41.80)	222 (52.0)	—[f]	206 (50.2)	—	82 (21.4)	—
Age (years)							
18-29	448 (36.72)	210 (49.2)	.13	180 (43.9)	.89	58 (15.1)	.58
≥30	772 (63.28)	217 (50.8)	—	230 (56.1)	—	325 (84.9)	—
Education level[g]							
≤High school diploma/General Education Development (GED) test	176 (14.42)	54 (12.6)	.96	31 (7.6)	.46	91 (23.8)	.76
Some post–high school education	387 (31.72)	128 (29.9)	—	98 (23.9)	—	161 (42.0)	—
4-year college degree	377 (30.90)	141 (33.0)	—	158 (38.5)	—	78 (20.4)	—
Some graduate education	277 (22.70)	102 (23.9)	—	123 (30.0)	—	52 (13.6)	—
City/metropolitan statistical area							
Atlanta	473 (38.77)	140 (32.8)	.98	145 (35.4)	.99	188 (49.1)	.98
Detroit	333 (27.29)	155 (36.3)	—	121 (29.5)	—	57 (14.9)	—
New York City	414 (33.93)	132 (30.9)	—	144 (35.1)	—	138 (36.0)	—
Baseline outcome variables							
STI[h] testing (3 months)	—	165 (38.4)	.18	254 (61.9)	.27	271 (71)	.09
HIV testing (3 months)	—	162 (37.9)	.40	259 (63.2)	.74	—	—
Current PrEP[i] use	—	22 (5.1)	.08	222 (54.1)	.21	—	—
PrEP adherence (≥25/30 days)[j]	—	11 of 22 (50.0)	.36	212 of 222 (95.5)	.99	—	—
Current ART[k] use	—	—	—	—	—	368 (96.1)	.99
ART adherence (≥25/30 days)[j]	—	—	—	—	—	335 of 368 (91.0)	.06
Unprotected anal sex (3 months)[l]	—	392 (91.8)	.37	51 (12.4)	.73	98 (25.6)	.01
Engagement in care (3 months)[m]	—	45 (10.5)	.10	196 (47.8)	.42	251 (65.5)	.06

[a]MSM: men who have sex with men.

[b]High risk (condomless/PrEP-less anal sex, past 3 months).

[c]Low risk (no condomless/PrEP-less anal sex, past 3 months).

[d]*P* value for chi-square testing (difference between intervention and control arms).

[e]One HIV-positive participant did not report race/ethnicity.

[f]Not applicable.

[g]Two high-risk HIV-negative and one HIV-positive participants did not report education level.

[h]STI: sexually transmitted infection.

[i]PrEP: preexposure prophylaxis.

[j]Among current users.

[k]ART: antiretroviral therapy.

[l]Condomless and PrEP-less anal sex for HIV-negative and condomless/detectable viral load for HIV-positive users.

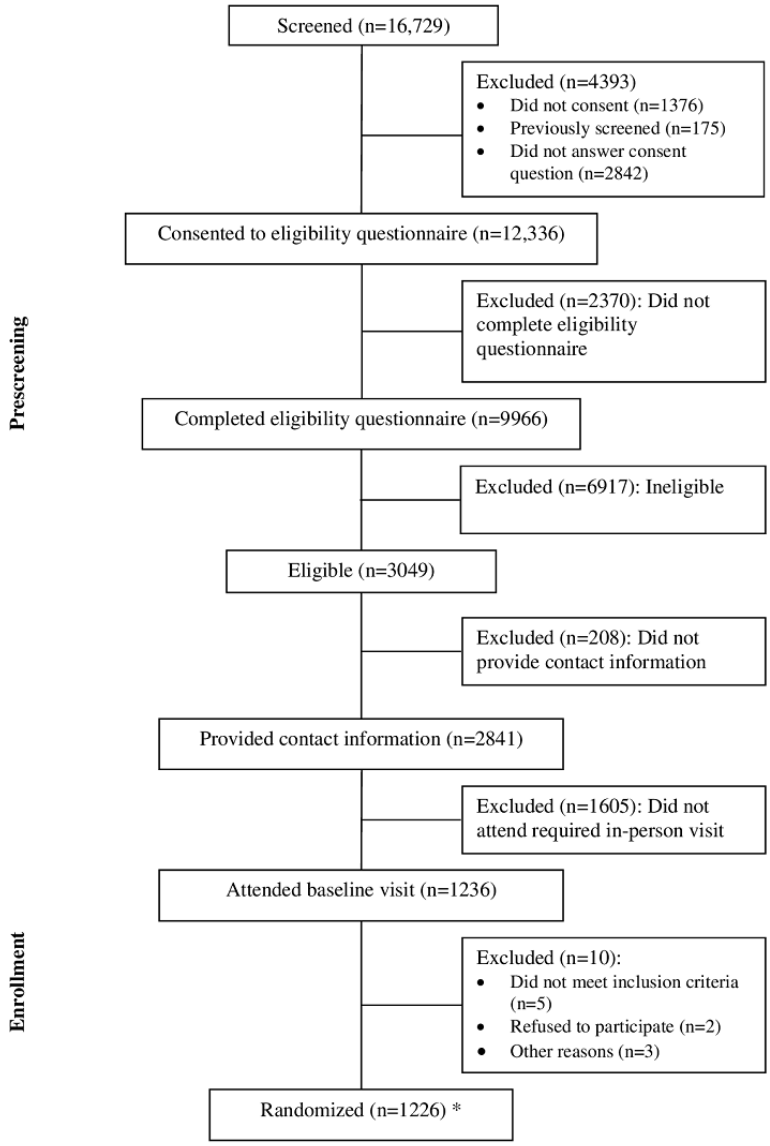
[m]Engagement in PrEP or ART care.

Results

Recruitment and Randomization

From January 24, 2018, to November 2, 2018, a total of 9966 GBMSM were assessed for eligibility for the randomized controlled trial; reasons for lack of eligibility are shown in Figure 1. Specifically, 2841 (28.51%) of 9966 GBMSM screened were eligible and provided contact information, of which 1236 (43.51%) attended a baseline study visit, 1226 (99.19%) consented and were randomized to intervention (n=611, 49.84%) and control (n=615, 50.16%) arms (Figure 1).

Figure 1. CONSORT diagram for the M-cubed randomized controlled trial among GBMSM (screening through randomization). CONSORT: Consolidated Standards of Reporting Trials; GBMSM: gay, bisexual, and other men who have sex with men; M-cubed: Mobile Messaging for Men.



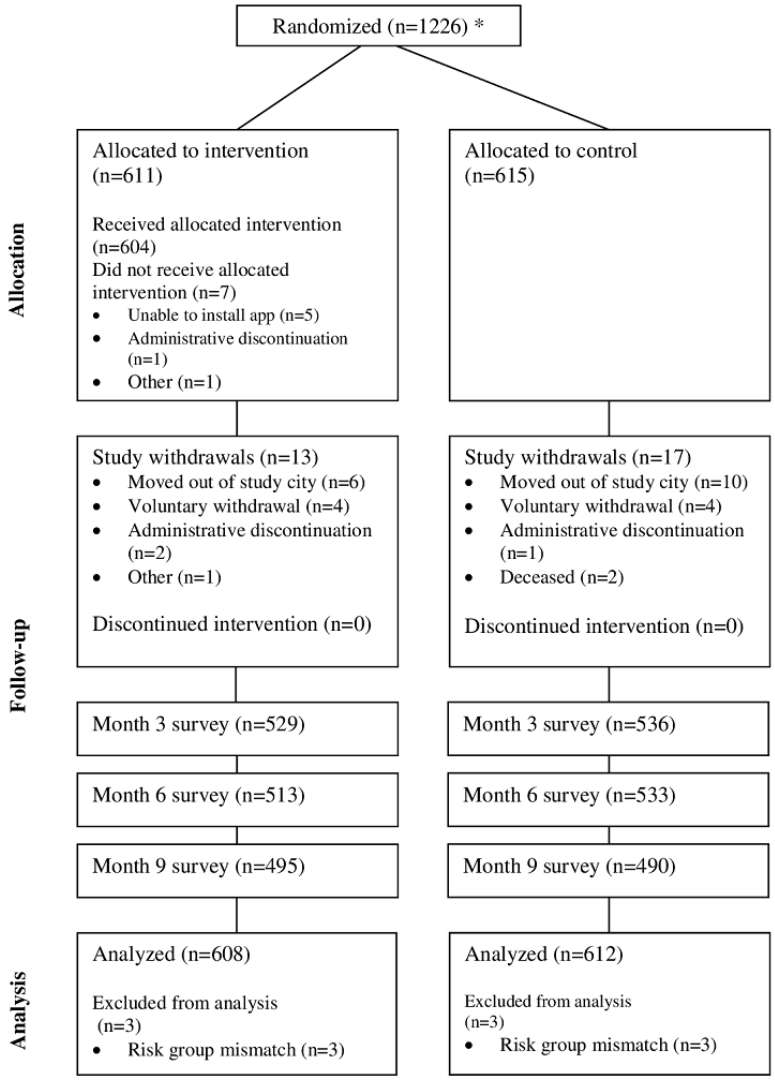
* Includes 388 in HIV-positive group, 410 in low-risk HIV-negative group, 428 in high-risk HIV-negative group

Sample Characteristics

The analytic sample comprised 1220 GBMSM enrolled in the study, including 608 (49.84%) randomized to the intervention arm and 612 (50.16%) randomized to the control arm (Figure 2). There was no failure of randomization within the 3 risk groups, except for unprotected anal intercourse among LWH GBMSM (Table 2). Most participants were GBMSM of color (n=378 [30.98%] Black/African American, n=183 [15%] Hispanic/Latino, n=148 [12.13%] other/mixed), were aged ≥30 years (772/1220 [63.28%]), and had at least a 4-year college education (654/1220 [53.61%]). Other indicators of risk and prevention services did not differ between control and intervention arms (data not shown). Self-reported levels of use of prevention and care services were often less than recommended by the Centers for Disease Control and Prevention

(CDC). High-risk HIV-negative GBMSM reported moderate levels of STI (165/427, 38.6%) and HIV (162/427, 37.9%) testing in the 3 months before baseline; only 22 of 427 (5.1%) reported using PrEP, and adherence to PrEP was low (11/22, 50%, reported ≥25/30 pills in the prior 30 days). A high proportion (392/427, 91.8%) reported sex not protected by PrEP or condoms in the past 3 months. Among low-risk HIV-negative GBMSM, baseline use of STI (254/410, 61.9%) and HIV (259/410, 63.2%) testing was higher; in addition, 222 of 410 (54.1%) reported being on PrEP, with high (212/222, 95.5%) adherence. Many HIV-negative GBMSM were designated as low risk because they were on PrEP; routine PrEP care includes HIV and STI testing. Among LWH GBMSM, most (271/383, 70.8%) reported recent STI testing, being on antiretroviral therapy (ART; 368/383, 96.1%) and high ART adherence (335/368, 91%). About 1 in 4 reported anal sex not protected by condoms in the past 3 months (98/383, 25.6%), and most (251/383, 65.5%) reported engaging in HIV care during the prior quarter.

Figure 2. CONSORT diagram for a randomized controlled trial of the M-cubed intervention for GBMSM (randomization through analysis). CONSORT: Consolidated Standards of Reporting Trials; GBMSM: gay, bisexual, and other men who have sex with men; M-cubed: Mobile Messaging for Men.



* Includes 388 in HIV-positive group, 410 in low-risk HIV-negative group, 428 in high-risk HIV-negative group

Retention was 1065 of 1226 (86.87%) at 3 months, 1046 of 1226 (85.32%) at 6 months, and 985 of 1226 (80.34%) at 9 months. Factors associated with retention were consistent across time points: retention was higher among participants who were White non-Hispanic, had more education, were employed full-time, and were not homeless. At the 9-month follow-up, higher retention was also associated with gay/homosexual identity (vs bisexual or other identities) and online recruitment (vs venue-based recruitment).

Intervention Efficacy

Intervention efficacy was analyzed by HIV status/risk group (Table 4). Among high-risk HIV-negative GBMSM, the odds of HIV testing (Figure 3) were higher at the immediate postintervention time point and the odds of current PrEP use

(Figure 4) were higher at the 3-month postintervention time point, but not at 6 or 9 months postintervention (Table 4). Of the high-risk GBMSM in the intervention arm who tested for HIV in the immediate postintervention period (n=126), 16 (12.7%) used an at-home HIV self-test kit provided through the app. As shown in Figure 4, although the initial prevalence of PrEP use was nonsignificantly lower in the intervention group at baseline, PrEP use rose from 3% at baseline to 15% by the 3-month follow-up assessment and remained at 15% through the 6-month follow-up assessment. Although odds ratio (OR) estimates comparing intervention to control for PrEP adherence were >1.0 at all postintervention time points for high-risk GBMSM, they were not statistically significant, given the small number of high-risk HIV-negative GBMSM taking PrEP.

For low-risk HIV-negative GBMSM, there were no changes in study outcomes associated with assignment to the intervention. For LWH GBMSM, STI testing was significantly lower in the immediate posttest period in the intervention group compared to the control group.

Table 4. Modeled behavioral measures at baseline, immediate posttest, and 3- and 6-month postintervention follow-up assessments for intervention efficacy of the Mobile Messaging for Men mobile app among MSM[a] by HIV status/risk group (N=1220).

Behavioral measure	High-risk HIV-negative MSM (n=427)			Low-risk HIV-negative MSM (n=410)			HIV-positive MSM (n=383)		
	Intervention group, M%[b] (95% CI)	Control group, M% (95% CI)	aOR[c] (95% CI)	Intervention group, M% (95% CI)	Control group, M% (95% CI)	aOR (95% CI)	Intervention group, M% (95% CI)	Control group, M% (95% CI)	aOR (95% CI)
STI[d] testing (3 months)									
Baseline	35 (28-43)	41 (34-49)	—[e]	58 (50-66)	64 (55-71)	—	76 (68-82)	68 (59-75)	—
Immediate posttest	42 (35-50)	47 (39-55)	1.07 (0.61-1.87)	61 (53-69)	56 (48-65)	1.53 (0.87-2.70)	65 (56-73)	72 (63-79)	*0.49 (0.26-0.94)*[f]
3 months postintervention	39 (31-47)	45 (37-53)	1.01 (0.58-1.76)	63 (54-71)	56 (48-64)	1.65 (0.97-2.81)	72 (63-79)	59 (50-68)	1.19 (0.63-2.24)
6 months postintervention	37 (30-45)	42 (34-50)	1.06 (0.64-1.74)	57 (48-66)	58 (49-66)	1.22 (0.77-1.94)	71 (62-78)	63 (54-72)	0.96 (0.53-1.73)
HIV testing (3 months)									
Baseline	48 (39-57)	52 (44-60)	—	73 (65-80)	71 (63-78)	—	—	—	—
Immediate posttest	63 (55-71)	51 (42-59)	*2.02 (1.11-3.66)*[f]	75 (67-81)	65 (57-73)	1.41 (0.75-2.64)	—	—	—
3 months postintervention	47 (39-55)	52 (44-60)	0.98 (0.55-1.75)	69 (61-77)	66 (58-73)	1.03 (0.57-1.86)	—	—	—
6 months postintervention	53 (45-61)	50 (42-58)	1.34 (0.80-2.25)	68 (60-76)	68 (59-75)	0.90 (0.54-1.52)	—	—	—
Current PrEP[g] use									
Baseline	3 (2-7)	7 (4-11)	—	55 (45-64)	49 (39-58)	—	—	—	—
Immediate posttest	9 (6-14)	15 (11-21)	1.26 (0.48-3.32)	52 (42-62)	49 (39-58)	0.89 (0.60-1.31)	—	—	—
3 months postintervention	15 (10-21)	14 (10-20)	*2.41 (1.00-5.76)*[f]	52 (42-61)	48 (39-59)	0.88 (0.63-1.23)	—	—	—
6 months postintervention	15 (11-21)	19 (14-26)	1.67 (0.81-3.47)	51 (42-61)	49 (39-58)	0.85 (0.66-1.10)	—	—	—
PrEP adherence (≥25/30 days)									
Baseline	35 (10-71)	62 (37-82)	—	95 (89-98)	95 (88-98)	—	—	—	—
Immediate posttest	78 (54-92)	80 (61-91)	2.72 (0.26-28.15)	91 (82-95)	84 (74-91)	1.66 (0.39-7.02)	—	—	—
3 months postintervention	84 (64-94)	74 (55-87)	5.47 (0.58-51.75)	92 (84-96)	85 (75-91)	1.84 (0.43-7.80)	—	—	—
6 months postintervention	85 (64-94)	92 (77-98)	1.38 (0.13-14.28)	91 (82-95)	92 (83-96)	0.88 (0.21-3.66)	—	—	—
Current ART[h] use									
Baseline	—	—	—	—	—	—	98 (94-99)	98 (95-99)	—
Immediate posttest	—	—	—	—	—	—	97 (92-77)	98 (94-99)	0.83 (0.17-4.15)

Behavioral measure	High-risk HIV-negative MSM (n=427)			Low-risk HIV-negative MSM (n=410)			HIV-positive MSM (n=383)		
	Intervention group, M%[b] (95% CI)	Control group, M% (95% CI)	aOR[c] (95% CI)	Intervention group, M% (95% CI)	Control group, M% (95% CI)	aOR (95% CI)	Intervention group, M% (95% CI)	Control group, M% (95% CI)	aOR (95% CI)
3 months postintervention	—	—	—	—	—	—	98 (94-99)	96 (92-98)	2.43 (0.56-10.55)
6 months postintervention	—	—	—	—	—	—	96 (91-98)	96 (91-98)	1.11 (0.38-3.30)
ART adherence (≥25/30 days)									
Baseline	—	—	—	—	—	—	96 (92-98)	91 (85-95)	—
Immediate posttest	—	—	—	—	—	—	93 (87-96)	91 (85-95)	0.58 (0.18-1.85)
3 months postintervention	—	—	—	—	—	—	93 (87-96)	92 (85-95)	0.40 (0.13-1.18)
6 months postintervention	—	—	—	—	—	—	92 (86-96)	92 (86-96)	0.43 (0.16-1.14)
Unprotected anal sex/all partners (3 months)[i]									
Baseline	94 (90-97)	92 (87-95)	—	14 (9-20)	15 (10-21)	—	21 (16-28)	34 (27-42)	—
Immediate posttest	78 (71-84)	73 (65-80)	0.93 (0.39-2.24)	23 (16-31)	28 (21-37)	0.79 (0.38-1.64)	20 (14-27)	20 (14-28)	1.93 (0.92-4.06)
3 months postintervention	72 (64-79)	76 (67-82)	0.58 (0.25-1.35)	22 (15-30)	29 (22-38)	0.75 (0.38-1.49)	15 (10-22)	19 (14-27)	1.46 (0.68-3.15)
6 months postintervention	70 (61-77)	66 (58-74)	0.82 (0.40-1.67)	27 (20-35)	26 (19-35)	1.11 (0.62-2.01)	18 (12-25)	18 (12-25)	1.98 (0.96-4.07)
Engagement in PrEP/ART care (3 months)									
Baseline	8 (5-13)	13 (9-19)	—	49 (40-59)	46 (37-55)	—	70 (63-77)	61 (53-68)	—
Immediate posttest	16 (11-22)	18 (13-25)	1.38 (0.62-3.04)	46 (37-56)	45 (36-55)	0.91 (0.54-1.55)	61 (53-69)	58 (50-66)	0.75 (0.40-1.39)
3 months postintervention	17 (12-24)	15 (10-22)	1.90 (0.89-4.04)	47 (37-57)	44 (35-54)	0.97 (0.60-1.57)	68 (60-76)	56 (48-64)	1.11 (0.60-2.04)
6 months postintervention	17 (12-24)	20 (14-27)	1.37 (0.73-2.56)	42 (33-52)	43 (34-53)	0.82 (0.55-1.22)	64 (55-72)	55 (47-63)	0.95 (0.54-1.65)

[a]MSM: men who have sex with men.

[b]M%, model-based mean probability of reporting behavior.

[c]aOR: adjusted odds ratio (for intervention vs control change from baseline; regression models include independent variables of the intervention group [intervention, control], survey time point [baseline, immediate posttest, 3- and 6-month postintervention follow-up], and their interaction).

[d]STI: sexually transmitted infection.

[e]Not applicable.

[f]$P<.05$.

[g]PrEP: preexposure prophylaxis.

[h]ART: antiretroviral therapy.

[i]Condomless/PrEP-less anal sex for HIV-negative MSM and condomless/detectable viral load for HIV-positive MSM.

Figure 3. HIV testing of high-risk GBMSM, reported by study time point and randomized allocation, in Atlanta, Detroit, and New York City (2018-2019). GBMSM: gay, bisexual, and other men who have sex with men.

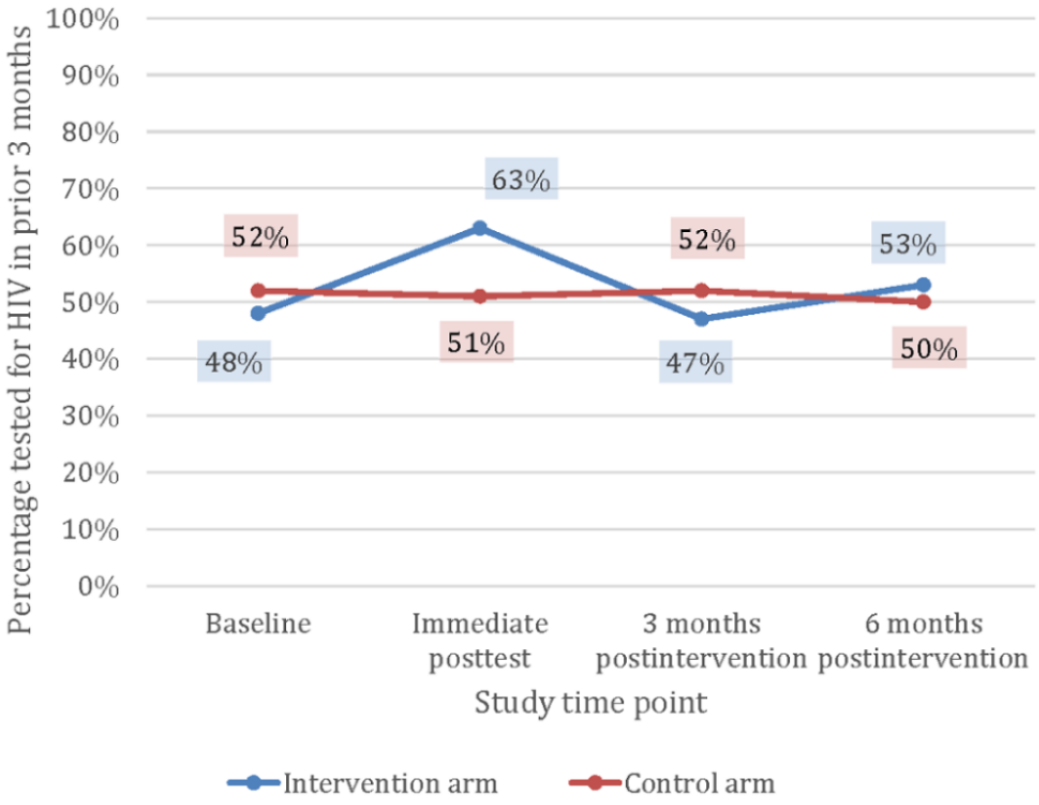
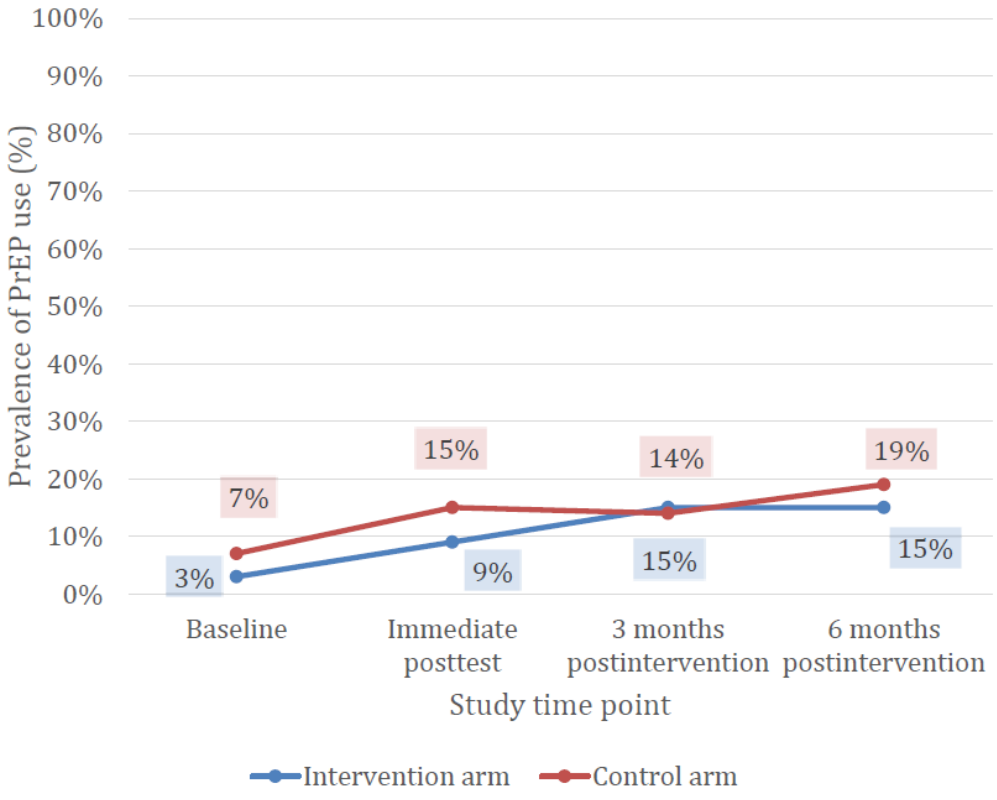


Figure 4. PrEP use among GBMSM who reported anal intercourse not protected by condoms or PrEP at baseline, reported by study time point and randomized allocation, in Atlanta, Detroit, and New York City (2018-2019). GBMSM: gay, bisexual, and other men who have sex with men; PrEP: preexposure prophylaxis.



Posthoc Analyses of Intervention Effects by Key Demographic Variables and Time Spent on App and STI Testing

Given the potential importance of race/ethnicity, age group, and education level as potential moderators of intervention effects, we conducted stratified analyses of the outcome variables for each of the 3 HIV status/risk groups for the following subgroups: non-Hispanic White versus others; age 18-29 versus ≥30 years; and <4-year college degree versus 4-year degree or more. No meaningful difference in outcomes was identified among stratum-specific estimates. Further, a sensitivity analysis assessed the time spent on the mobile app over the 3-month intervention (control vs ≥30 minutes of intervention use) for the 3 HIV status/risk groups. No meaningful or significant difference in outcomes was identified.

In a posthoc analysis, the intention to test for STIs (ie, ordering an STI test kit or reporting having had an STI test) was higher in all HIV-negative participants (low-risk HIV-negative: aOR 3.1, 95% CI 1.7-5.6; high-risk HIV-negative: aOR 3.5, 95% CI 2.0-6.2; all HIV-negative regardless of risk group: aOR 3.4, 95% CI 2.3-5.1).

Discussion

Principal Results

We tested an app-based intervention to provide messages to promote HIV prevention and care and provide access to core HIV and STI prevention services to GBMSM. The results showed significant increases in HIV screening during the period of app use among high-risk HIV-negative GBMSM. The results also indicated increased prevalence of PrEP use among high-risk HIV-negative GBMSM in the 3-month period after app use. We identified no protective changes in study outcomes for low-risk HIV-negative GBMSM or LWH GBMSM. Given the high baseline PrEP use and adherence among low-risk HIV-negative GBMSM who did not report recent anal sex not protected by condoms and high levels of ART use and the high baseline adherence among LWH GBMSM, there was little room for improvement statistically among these groups on key outcome variables. We recognize, however, that the vulnerability of GBMSM may change over time (ie, seasons of risk [34]), so the utility of the app and its messaging might change over time for individual men.

We developed our intervention in light of 3 principles. First, we worked from the paradigm of integrated HIV prevention and care continua. Although the HIV status was determined at the beginning of the trial, when the app is used in practice, HIV testing will be the starting point for both continua. Thus, the intervention could serve GBMSM who do not know their HIV status [35]. Second, we sought to develop an intervention that could be offered to all GBMSM rather than to demographic or risk subsets of men. Third, we sought to develop an intervention that could be scaled to provide intervention content to large numbers of men without the need for a large staff of interventionists. A free, downloadable app that disseminates one 2-sentence message every other day and a 1-minute video weekly might be easily incorporated into a user's life without requiring the time and travel often required of more traditional, in-person interventions. Further, mobile app interventions are a potentially low-cost strategy from the perspective of health departments and community-based organizations (CBOs) that may choose to implement them in their jurisdictions.

We view our mixed efficacy outcomes from several perspectives. The fact that exposure to the app was associated with a doubling of HIV-testing behaviors and of prevalence of PrEP use among high-risk HIV-negative GBMSM suggests that the messages and app services could be a scalable, effective resource for GBMSM who engage in anal intercourse that is not protected by condoms or PrEP. Large national surveys of GBMSM suggest that 47%-53% of HIV-negative GBMSM would fall into this category [36,37]; considering the number of high-risk HIV-negative GBMSM in the United States [38], the potential user base for these services is about 1,800,000-2,000,000 US GBMSM.

The effect sizes for uptake of HIV screening and PrEP use were both about 2. However, the baseline levels of HIV screening and PrEP use were quite different. For high-risk GBMSM, the CDC recommends HIV testing at least annually, and in the intervention arm, nearly two-thirds reported HIV testing in the immediate posttest assessment period (tested in the prior 3 months). However, despite the same relative increase in PrEP use, only 1 in 6 (17%) high-risk GBMSM was taking PrEP at the 3-month postintervention time point. This suggests that additional and more intensive intervention services will likely be required for high-risk PrEP-eligible men who do not initiate PrEP within 6 months of starting the use of the Mobile Messaging for Men (M-cubed) app.

We also recognize that increases in PrEP usage and HIV testing associated with allocation to the app were short lived. The timing of the effects makes sense: the support for HIV testing provided during the period of app use included ordering at-home kits and navigating to testing locations; the significant increase in testing was observed during the period when men had access to the app, and about 1 in 7 high-risk men who tested for HIV used a mailed-out HIV self-test kit provided through the study. For PrEP, after the decision to start PrEP, there is a long process of making an appointment, undergoing lab tests, filling a prescription, and starting the medication; this outcome was higher in the 3 months after the period of app use. In practice, exposure to the app would not be limited to 3 months, and it is a separate question as to whether changes in behavior would be sustained over time with ongoing exposure to the app. Such a question could be answered in an implementation study.

Our approach was developed with an eye toward being able to reach large numbers of GBMSM with a basic package of prevention services. Given that there was efficacy for some outcomes, there are important questions about how the intervention might be made available to GBMSM who would be likely to benefit from it [39]. Most HIV prevention services are currently provided and funded through CBOs and health departments. Although some CBOs and health departments might have the technical capacity to support app-based prevention services, this capacity is likely quite limited. A currently underway randomized study of an efficacious eHealth

intervention (Keep It Up [40]) will examine the outcomes of centralized versus decentralized implementation approaches, and the results of the study may inform how the M-cubed intervention could be most effectively implemented [41].

Limitations

Our study was subject to several limitations. Although we used recruitment quotas to guide recruitment and achieved a balanced sample with respect to the 3 HIV status/risk groups, race/ethnicity, and recruitment venues, our sample was skewed to men with higher education. These selection biases threaten the external generalizability of the study; it is possible that the intervention might have had salutary effects for low-risk HIV-negative GBMSM or LWH GBMSM with lower education. However, although not powered for secondary comparisons, our post hoc stratified analyses did not suggest differential effects by educational attainment. Second, PrEP use in our sample was high at baseline (about 29% among all HIV-negative GBMSM), which is unique in large samples of GBMSM to date. For example, current PrEP use was reported by less than 20% of respondents to the American Men's Internet Survey (a large, national sample of US GBMSM) in 2017 [10] and by 25% of GBMSM who participated in 23 cities in the 2017 National HIV Behavioral Surveillance study [37]. Over half of GBMSM in our study who were not on PrEP reported HIV testing in the 3 months before enrollment; this is similar to the proportion who tested for HIV in the 12 months before interviews in the 2016 AMIS survey [11]. Third, outcomes were self-reported and so might be subject to misclassification because of social desirability bias [42]. Fourth, our sample was restricted to GBMSM recruited online; this might introduce selection bias, in that men recruited through physical venues, such as sex venues, might have higher or lower levels of risk. Lastly, our sample was limited to GBMSM in 3 urban areas of the United States who reported sex with a man in the past year; patterns of risk behaviors and use of prevention services are known to be different in other urban areas or among rural GBMSM [11,21]. The level of ART use in our LWH MSM was high, offering limited opportunity for improvement. Future studies of the app could condition enrollment on the need for support for ART initiation or maintenance for LWH MSM.

The next steps for the study include secondary analysis of extensive app usage data or *paradata* [43] in terms of understanding the ratings of each message and the time spent using various optional aspects of the app descriptively [44]. We will also conduct further assessments of the potential mediation of outcome effects by app usage to help direct future prevention research and program implementation of mobile apps. Based on our data, the app offered benefits to GBMSM who reported recent anal sex not protected by condoms or PrEP but not to HIV-negative men who did not report sex not protected by condoms of PrEP or LWH men. Therefore, we will consider how the app can be updated and implemented [39,45] for GBMSM engaging in sex not protected by condoms or PrEP and for LWH MSM. Because we found that the app is a successful way to increase the intention to test for STIs and to distribute at-home STI self-test kits, the kits were often not returned and additional research is needed to understand the barriers to returning STI specimen collection kits. If this barrier can be overcome, the M-cubed app might also effectively support increased STI testing for some groups of GBMSM.

Conclusion

We sought to produce and evaluate a mobile app with prevention messaging and services to support comprehensive HIV prevention and care with low levels of prevention staff interaction for all GBMSM, regardless of HIV status or current risk behaviors. Such a serostatus-neutral approach, if implemented successfully, would offer advantages in terms of cost efficiency and minimization of segmentation of GBMSM by serostatus. Future work with the M-cubed app should focus on implementing the app for GBMSM who are likely to benefit from it based on their current behaviors and evaluating methods of implementation of the app through public or publicly supported prevention providers [39].

Acknowledgments

This study was funded through the US Centers for Disease Control and Prevention (CDC) to Emory University (cooperative agreement no. U01PS004977) and the US National Institutes of Health (P30AI050409).

CDC disclaimer: The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the CDC.

Authors' Contributions

Conceptualization, formal analysis, funding acquisition, investigation, methodology, supervision, writing (original draft), and writing (review and editing) were performed by PSS. Conceptualization, funding acquisition, investigation, supervision, and writing (review and editing) were performed by RS and SH. Formal analysis, methodology, supervision, writing (review and editing) were conducted by CCM. RZ performed funding acquisition, investigation, supervision, and writing (review and editing); JAB and KH performed funding acquisition, methodology, and writing (review and editing); MAC conducted conceptualization, funding acquisition, investigation, writing (review and editing); DG performed conceptualization, methodology, supervision, and writing (review and editing); SM performed formal analysis, methodology, supervision, and writing (review and editing); and MJD, EJO, and SW conducted investigation, methodology, and writing (review and editing). EQR performed investigation, methodology, supervision, and writing (review and editing); ER and AJS performed conceptualization, funding acquisition,

methodology, and writing (review and editing); and GM conducted conceptualization, formal analysis, investigation, methodology, supervision, writing (original draft), and writing (review and editing). All authors read and approved the final manuscript.
The authors have no conflicts of interest to disclose. In this cooperative agreement, the US Centers for Disease Control and Prevention (CDC) scientists participated in the design of this study and will play collaborative roles during its execution, analyses, and interpretation and dissemination of the data.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 736 KB - jmir_v24i2e34574_app1.pdf]

References

1. Fauci AS, Redfield RR, Sigounas G, Weahkee MD, Giroir BP. Ending the HIV epidemic: a plan for the United States. JAMA 2019 Mar 05;321(9):844-845. [doi: 10.1001/jama.2019.1343] [Medline: 30730529]
2. Sullivan P, Woodyatt C, Koski C, Pembleton E, McGuinness P, Taussig J, et al. A data visualization and dissemination resource to support HIV prevention and care at the local level: analysis and uses of the AIDSVu public data resource. J Med Internet Res 2020 Oct 23;22(10):e23173 [FREE Full text] [doi: 10.2196/23173] [Medline: 33095177]
3. Centers for Disease Control and Prevention. Diagnoses of HIV Infection in the United States and Dependent Areas, 2018. URL: https://www.cdc.gov/hiv/library/reports/hiv-surveillance/vol-31/index.html [accessed 2022-01-19]
4. Centers for Disease Control and Prevention. Complete Listing of Risk Reduction Evidence-based Behavioral Interventions Internet. URL: https://www.cdc.gov/hiv/research/interventionresearch/compendium/rr/complete.html [accessed 2022-01-19]
5. Maloney KM, Bratcher A, Wilkerson R, Sullivan PS. Electronic and other new media technology interventions for HIV care and prevention: a systematic review. J Int AIDS Soc 2020 Jan;23(1):e25439 [FREE Full text] [doi: 10.1002/jia2.25439] [Medline: 31909896]
6. Hightow-Weidman LB, Muessig K, Rosenberg E, Sanchez T, LeGrand S, Gravens L, et al. University of North Carolina/Emory Center for Innovative Technology (iTech) for addressing the HIV epidemic among adolescents and young adults in the United States: protocol and rationale for center development. JMIR Res Protoc 2018 Aug 03;7(8):e10365 [FREE Full text] [doi: 10.2196/10365] [Medline: 30076126]
7. Sullivan PS, Peterson J, Rosenberg ES, Kelley CF, Cooper H, Vaughan A, et al. Understanding racial HIV/STI disparities in black and white men who have sex with men: a multilevel approach. PLoS One 2014 Mar 7;9(3):e90514 [FREE Full text] [doi: 10.1371/journal.pone.0090514] [Medline: 24608176]
8. Hightow-Weidman LB, Magnus M, Beauchamp G, Hurt CB, Shoptaw S, Emel L, et al. Incidence and correlates of sexually transmitted infections among black men who have sex with men participating in the HIV Prevention Trials Network 073 Preexposure Prophylaxis Study. Clin Infect Dis 2019 Oct 15;69(9):1597-1604 [FREE Full text] [doi: 10.1093/cid/ciy1141] [Medline: 30615169]
9. Sullivan PS, Satcher Johnson A, Pembleton ES, Stephenson R, Justice AC, Althoff KN, et al. Epidemiology of HIV in the USA: epidemic burden, inequities, contexts, and responses. Lancet 2021 Mar;397(10279):1095-1106. [doi: 10.1016/s0140-6736(21)00395-0]
10. Sullivan PS, Sanchez TH, Zlotorzynska M, Chandler CJ, Sineath R, Kahle E, et al. National trends in HIV pre-exposure prophylaxis awareness, willingness and use among United States men who have sex with men recruited online, 2013 through 2017. J Int AIDS Soc 2020 Mar;23(3):e25461 [FREE Full text] [doi: 10.1002/jia2.25461] [Medline: 32153119]
11. Sanchez TH, Zlotorzynska M, Sineath RC, Kahle E, Tregear S, Sullivan PS. National trends in sexual behavior, substance use and HIV testing among United States men who have sex with men recruited online, 2013 through 2017. AIDS Behav 2018 Aug 12;22(8):2413-2425. [doi: 10.1007/s10461-018-2168-4] [Medline: 29948340]
12. Sullivan PS, Carballo-Diéguez A, Coates T, Goodreau SM, McGowan I, Sanders EJ, et al. Successes and challenges of HIV prevention in men who have sex with men. Lancet 2012 Jul;380(9839):388-399. [doi: 10.1016/s0140-6736(12)60955-6]
13. Jenness SM, Goodreau SM, Rosenberg E, Beylerian EN, Hoover KW, Smith DK, et al. Impact of the Centers for Disease Control's HIV preexposure prophylaxis guidelines for men who have sex with men in the United States. J Infect Dis 2016 Dec 15;214(12):1800-1807 [FREE Full text] [doi: 10.1093/infdis/jiw223] [Medline: 27418048]
14. Brookmeyer R, Boren D, Baral SD, Bekker LG, Phaswana-Mafuya N, Beyrer C, et al. Combination HIV prevention among MSM in South Africa: results from agent-based modeling. PLoS One 2014 Nov 14;9(11):e112668 [FREE Full text] [doi: 10.1371/journal.pone.0112668] [Medline: 25398143]
15. Hightow-Weidman LB, Muessig KE, Bauermeister J, Zhang C, LeGrand S. Youth, technology, and HIV: recent advances and future directions. Curr HIV/AIDS Rep 2015 Dec;12(4):500-515 [FREE Full text] [doi: 10.1007/s11904-015-0280-x] [Medline: 26385582]

16. Muessig KE, Nekkanti M, Bauermeister J, Bull S, Hightow-Weidman LB. A systematic review of recent smartphone, internet and Web 2.0 interventions to address the HIV continuum of care. Curr HIV/AIDS Rep 2015 Mar 28;12(1):173-190 [FREE Full text] [doi: 10.1007/s11904-014-0239-3] [Medline: 25626718]
17. Muessig KE, Pike EC, Legrand S, Hightow-Weidman LB. Mobile phone applications for the care and prevention of HIV and other sexually transmitted diseases: a review. J Med Internet Res 2013 Jan 04;15(1):e1 [FREE Full text] [doi: 10.2196/jmir.2301] [Medline: 23291245]
18. Sullivan P, Grey J, Simon Rosser BR. Emerging technologies for HIV prevention for MSM: what we have learned, and ways forward. J Acquir Immune Defic Syndr 2013 Jun 01;63 Suppl 1:S102-S107 [FREE Full text] [doi: 10.1097/QAI.0b013e3182949e85] [Medline: 23673879]
19. Sullivan PS, Jones J, Kishore N, Stephenson R. The roles of technology in primary HIV prevention for men who have sex with men. Curr HIV/AIDS Rep 2015 Dec 30;12(4):481-488. [doi: 10.1007/s11904-015-0293-5] [Medline: 26519083]
20. Muessig KE, Pike EC, Fowler B, LeGrand S, Parsons JT, Bull SS, et al. Putting prevention in their pockets: developing mobile phone-based HIV interventions for black men who have sex with men. AIDS Patient Care STDS 2013 Apr;27(4):211-222 [FREE Full text] [doi: 10.1089/apc.2012.0404] [Medline: 23565925]
21. McKenney J, Sullivan PS, Bowles KE, Oraka E, Sanchez TH, DiNenno E. HIV risk behaviors and utilization of prevention services, urban and rural men who have sex with men in the United States: results from a national online survey. AIDS Behav 2018 Jul 6;22(7):2127-2136. [doi: 10.1007/s10461-017-1912-5] [Medline: 28986669]
22. Khosropour CM, Lake JG, Sullivan PS. Are MSM willing to SMS for HIV prevention? J Health Commun 2014 Aug 26;19(1):57-66. [doi: 10.1080/10810730.2013.798373] [Medline: 23905653]
23. Compendium of Evidence-Based Interventions and Best Practices for HIV Prevention Internet. URL: http://www.cdc.gov/hiv/prevention/research/compendium/index.html [accessed 2022-01-19]
24. Fernandez MI, Harper GW, Hightow-Weidman LB, Kapogiannis BG, Mayer KH, Parsons JT, et al. Research priorities to end the adolescent HIV epidemic in the United States: viewpoint. JMIR Res Protoc 2021 Jan 04;10(1):e22279 [FREE Full text] [doi: 10.2196/22279] [Medline: 33393918]
25. Sullivan PS, Driggers R, Stekler JD, Siegler A, Goldenberg T, McDougal SJ, et al. Usability and acceptability of a mobile comprehensive HIV prevention app for men who have sex with men: a pilot study. JMIR Mhealth Uhealth 2017 Mar 09;5(3):e26 [FREE Full text] [doi: 10.2196/mhealth.7199] [Medline: 28279949]
26. Goldenberg T, McDougal SJ, Sullivan PS, Stekler JD, Stephenson R. Building a mobile HIV prevention app for men who have sex with men: an iterative and community-driven process. JMIR Public Health Surveill 2015 Nov 16;1(2):e18 [FREE Full text] [doi: 10.2196/publichealth.4449] [Medline: 27227136]
27. Goldenberg T, McDougal SJ, Sullivan PS, Stekler JD, Stephenson R. Preferences for a mobile HIV prevention app for men who have sex with men. JMIR Mhealth Uhealth 2014 Oct 29;2(4):e47 [FREE Full text] [doi: 10.2196/mhealth.3745] [Medline: 25355249]
28. Sullivan PS, Zahn RJ, Wiatrek S, Chandler CJ, Hirshfield S, Stephenson R, et al. HIV prevention via mobile messaging for men who have sex with men (M-cubed): protocol for a randomized controlled trial. JMIR Res Protoc 2019 Nov 15;8(11):e16439. [doi: 10.2196/16439]
29. Bauermeister JA, Golinkoff JM, Horvath KJ, Hightow-Weidman LB, Sullivan PS, Stephenson R. A multilevel tailored web app-based intervention for linking young men who have sex with men to quality care (get connected): protocol for a randomized controlled trial. JMIR Res Protoc 2018 Aug 02;7(8):e10444 [FREE Full text] [doi: 10.2196/10444] [Medline: 30072358]
30. Horvath K, Amico K, Erickson D, Ecklund A, Martinka A, DeWitt J, et al. Thrive with me: protocol for a randomized controlled trial to test a peer support intervention to improve antiretroviral therapy adherence among men who have sex with men. JMIR Res Protoc 2018 May 31;7(5):e10182 [FREE Full text] [doi: 10.2196/10182] [Medline: 29853437]
31. Kernan W, Viscoli C, Makuch R, Brass L, Horwitz R. Stratified randomization for clinical trials. J Clin Epidemiol 1999 Jan;52(1):19-26. [doi: 10.1016/s0895-4356(98)00138-3] [Medline: 9973070]
32. Sanchez TH, Sineath RC, Kahle EM, Tregear SJ, Sullivan PS. The Annual American Men's Internet Survey of behaviors of men who have sex with men in the United States: protocol and key indicators report 2013. JMIR Public Health Surveill 2015 Apr 17;1(1):e3 [FREE Full text] [doi: 10.2196/publichealth.4314] [Medline: 27227126]
33. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Ann Intern Med 2010 Jun 01;152(11):726-732 [FREE Full text] [doi: 10.7326/0003-4819-152-11-201006010-00232] [Medline: 20335313]
34. Elsesser SA, Oldenburg CE, Biello KB, Mimiaga MJ, Safren SA, Egan JE, et al. Seasons of risk: anticipated behavior on vacation and interest in episodic antiretroviral pre-exposure prophylaxis (PrEP) among a large national sample of U.S. men who have sex with men (MSM). AIDS Behav 2016 Jul 4;20(7):1400-1407 [FREE Full text] [doi: 10.1007/s10461-015-1238-0] [Medline: 26538056]
35. Myers JE, Braunstein SL, Xia Q, Scanlin K, Edelstein Z, Harriman G, et al. Redefining prevention and care: a status-neutral approach to HIV. Open Forum Infect Dis 2018 Jun;5(6):ofy097 [FREE Full text] [doi: 10.1093/ofid/ofy097] [Medline: 29977957]

36. Zlotorzynska M, Cantu C, Rai R, Sullivan P, Sanchez T. The Annual American Men's Internet Survey of behaviors of men who have sex with men in the United States: 2017 key indicators report. JMIR Public Health Surveill 2020 Apr 13;6(2):e16847 [FREE Full text] [doi: 10.2196/16847] [Medline: 32281937]
37. National HIV Behavioral Surveillance (NHBS) Study Group. HIV Infection Risk, Prevention, and Testing Behaviors Among Men Who Have Sex With Men? National HIV Behavioral Surveillance, 23 U.S. Cities, 2017. 2019 Feb. URL: https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-special-report-number-22.pdf [accessed 2022-01-19]
38. Rosenberg ES, Grey JA, Sanchez TH, Sullivan PS. Rates of prevalent HIV infection, prevalent diagnoses, and new diagnoses among men who have sex with men in US states, metropolitan statistical areas, and counties, 2012-2013. JMIR Public Health Surveill 2016 May 17;2(1):e22 [FREE Full text] [doi: 10.2196/publichealth.5684] [Medline: 27244769]
39. Li DH, Brown CH, Gallo C, Morgan E, Sullivan PS, Young SD, et al. Design considerations for implementing eHealth behavioral interventions for HIV prevention in evolving sociotechnical landscapes. Curr HIV/AIDS Rep 2019 Aug 27;16(4):335-348 [FREE Full text] [doi: 10.1007/s11904-019-00455-4] [Medline: 31250195]
40. Mustanski B, Parsons JT, Sullivan PS, Madkins K, Rosenberg E, Swann G. Biomedical and behavioral outcomes of Keep It Up! An eHealth HIV prevention program RCT. Am J Prev Med 2018 Aug;55(2):151-158 [FREE Full text] [doi: 10.1016/j.amepre.2018.04.026] [Medline: 29937115]
41. Mustanski B, Jones J, Macapagal K, Benbow N, Li D, Hendricks BC, et al. Keep It Up! 3.0: Study Protocol for a Type III Hybrid Implementation-Effectiveness Cluster-Randomized Trial Internet. 202 Nov 13. URL: https://www.researchsquare.com/article/rs-104488/v1 [accessed 2022-01-19]
42. Sackett DL. Bias in analytic research. J Chronic Dis 1979 Jan;32(1-2):51-63. [doi: 10.1016/0021-9681(79)90012-2]
43. Bauermeister J, Golinkoff J, Muessig K, Horvath K, Hightow-Weidman L. Addressing engagement in technology-based behavioural HIV interventions through paradata metrics. Curr Opin HIV AIDS 2017;12(5):446. [doi: 10.1097/coh.0000000000000396]
44. Hightow-Weidman LB, Bauermeister JA. Engagement in mHealth behavioral interventions for HIV prevention and care: making sense of the metrics. Mhealth 2020;6:7 [FREE Full text] [doi: 10.21037/mhealth.2019.10.01] [Medline: 32190618]
45. Sullivan PS, Hightow-Weidman L. Mobile apps for HIV prevention: how do they contribute to our epidemic response for adolescents and young adults? Mhealth 2021;7:36 [FREE Full text] [doi: 10.21037/mhealth-20-71] [Medline: 33898605]

Abbreviations

aOR: adjusted odds ratio
ART: antiretroviral therapy
CBO: community-based organization
CDC: Centers for Disease Control and Prevention
CONSORT: Consolidated Standards of Reporting Trials
GBMSM: gay, bisexual, and other men who have sex with men
LWH: living with HIV
M-cubed: Mobile Messaging for Men
mHealth: mobile health
MSM: men who have sex with men
nPEP: nonoccupational postexposure prophylaxis
OR: odds ratio
PrEP: preexposure prophylaxis
SMART: Study Management and Retention Tool
STI: sexually transmitted infection
YGBMSM: young gay, bisexual, and other men who have sex with men

Edited by A Mavragani; submitted 29.11.21; peer-reviewed by P Serrano; comments to author 20.12.21; revised version received 20.12.21; accepted 21.12.21; published 02.02.22.

Please cite as:
Sullivan PS, Stephenson R, Hirshfield S, Mehta CC, Zahn R, Bauermeister JA, Horvath K, Chiasson MA, Gelaude D, Mullin S, Downing Jr MJ, Olansky EJ, Wiatrek S, Rogers EQ, Rosenberg E, Siegler AJ, Mansergh G
Behavioral Efficacy of a Sexual Health Mobile App for Men Who Have Sex With Men: Randomized Controlled Trial of Mobile Messaging for Men
J Med Internet Res 2022;24(2):e34574
URL: https://www.jmir.org/2022/2/e34574
doi:10.2196/34574
PMID:35025755

Original Paper

Measuring Electronic Health Literacy: Development, Validation, and Test of Measurement Invariance of a Revised German Version of the eHealth Literacy Scale

Matthias Marsall[1,2], MSc; Gerrit Engelmann[1], MSc; Eva-Maria Skoda[1], PD, MD; Martin Teufel[1], Prof Dr; Alexander Bäuerle[1], PhD

[1]Clinic for Psychosomatic Medicine and Psychotherapy, LVR–University Hospital Essen, University of Duisburg-Essen, Essen, Germany
[2]Institute for Patient Safety, University Hospital Bonn, Bonn, Germany

Corresponding Author:
Matthias Marsall, MSc
Clinic for Psychosomatic Medicine and Psychotherapy
LVR–University Hospital Essen
University of Duisburg-Essen
Virchowstr. 174
Essen, 45147
Germany
Phone: 49 17678909441
Email: matthias.marsall@stud.uni-due.de

Abstract

Background: The World Wide Web has become an essential source of health information. Nevertheless, the amount and quality of information provided may lead to information overload. Therefore, people need certain skills to search for, identify, and evaluate information from the internet. In the context of health information, these competencies are summarized as the construct of eHealth literacy. Previous research has highlighted the relevance of eHealth literacy in terms of health-related outcomes. However, the existing instrument assessing eHealth literacy in the German language reveals methodological limitations regarding test development and validation. The development and validation of a revised scale for this important construct is highly relevant.

Objective: The objective of this study was the development and validation of a revised German eHealth literacy scale. In particular, this study aimed to focus on high methodological and psychometric standards to provide a valid and reliable instrument for measuring eHealth literacy in the German language.

Methods: Two internationally validated instruments were merged to cover a wide scope of the construct of eHealth literacy and create a revised eHealth literacy scale. Translation into the German language followed scientific guidelines and recommendations to ensure content validity. Data from German-speaking people (n=470) were collected in a convenience sample from October to November 2020. Validation was performed by factor analyses. Further, correlations were performed to examine convergent, discriminant, and criterion validity. Additionally, analyses of measurement invariance of gender, age, and educational level were conducted.

Results: Analyses revealed a 2-factorial model of eHealth literacy. By item-reduction, the 2 factors information seeking and information appraisal were measured with 8 items reaching acceptable-to-good model fits (comparative fit index [CFI]: 0.942, Tucker Lewis index [TLI]: 0.915, root mean square error of approximation [RMSEA]: 0.127, and standardized root mean square residual [SRMR]: 0.055). Convergent validity was comprehensively confirmed by significant correlations of information seeking and information appraisal with health literacy, internet confidence, and internet anxiety. Discriminant and criterion validity were examined by correlation analyses with various scales and could partly be confirmed. Scalar level of measurement invariance for gender (CFI: 0.932, TLI: 0.923, RMSEA: 0.122, and SRMR: 0.068) and educational level (CFI: 0.937, TLI: 0.934, RMSEA: 0.112, and SRMR: 0.063) were confirmed. Measurement invariance of age was rejected.

Conclusions: Following scientific guidelines for translation and test validation, we developed a revised German eHealth Literacy Scale (GR-eHEALS). Our factor analyses confirmed an acceptable-to-good model fit. Construct validation in terms of convergent, discriminant, and criterion validity could mainly be confirmed. Our findings provide evidence for measurement invariance of the instrument regarding gender and educational level. The newly revised GR-eHEALS questionnaire represents a valid instrument to measure the important health-related construct eHealth literacy.

(J Med Internet Res 2022;24(2):e28252) doi:10.2196/28252

KEYWORDS

eHealth; eHeals; health literacy; factor analysis; validation; measurement invariance; internet; health information

Introduction

Background

The concept of health literacy emerged in the 1990s as a competence to gather health information and use it to address health questions and problems [1]. Nutbeam [2] defined health literacy as "cognitive and social skills which determine the motivation and ability of individuals to gain access to, understand, and use information in ways which promote and maintain good health." In the following years, health literacy has turned out to be an important predictor for various health outcomes (eg, behavior of patients with diabetes mellitus or heart failure) [3,4]. The World Health Organization has declared health literacy as a key determinant of health and defined it as a Sustainable Development Goal [5].

With the rise of the internet as a source of information, the gathering of health information was no longer limited to professional or face-to-face health sources but was available from many different health topic websites [6]. With the increasing availability of health information on the internet, the number of people using this source for seeking health information rose as well [7,8]. However, sources on the internet contain inconsistent information as contributions are not by professionals only [9]. As a result, the amount and differences in quality of information provided on the internet may lead to health information overload [10]. For example, in 2020, COVID-19 became a global pandemic, and disease-related information, especially from the internet, grew exponentially, leading to an "infodemic" [11,12]. Not only is a large amount of information available, but a significant amount of it must be considered misinformation because the sources of the information must be classified questionable [13,14].

For the context of information from the internet, Norman and Skinner [15] applied the concept of health literacy to electronic health literacy (eHealth literacy). With the development of the eHealth Literacy Scale (eHEALS) questionnaire [16], the concept of eHealth literacy became measurable and emerged as a growing interest in psychological and medical health sciences. Systematic reviews have shown that eHEALS is associated with different health-related outcomes, but findings could not be consistently confirmed [17,18]. Associations of eHealth literacy with different health outcomes have been found, such as health intentions [19], acquiring health knowledge [20-23], and health prevention behavior [21,24,25]. Furthermore, research showed associations between eHealth literacy and healthy behaviors like exercise behavior, balanced nutrition, and regular breakfast [26,27]. In the context of COVID-19, associations of eHealth literacy and lower psychological symptoms [28] and higher prevention behaviors [29] could be confirmed. To sum up, research indicates that eHealth literacy is associated with prevention behaviors, the acquisition of knowledge, and people's ability to cope with diseases, which confirms eHealth literacy as an important construct in examining people's health behavior.

To cope with information overload and use the information from the internet, Norman and Skinner [15] proposed a set of different competencies: skills to read, identify, and understand different information to distinguish helpful from less helpful or even false or harmful information. These competencies represent a sequential process of handling available information. In the first step, basic cognitive skills are needed to search for information regarding a certain topic. In a subsequent cognitive process, information available must be distinguished as helpful or less helpful in order to answer specific questions. These steps represent an elaborated cognitive information process rather than a heuristic one. The distinction of cognitive processes was formerly described within dual-process theories in psychological literature and confirmed in multiple studies [30-32]. Dual-process theories distinguish between fast cognitive processes, which describe heuristic and holistic approaches representing intuitive, implicit cognitions, and slow cognitive processes, which are analytic and rule-based and focus on explicit learning [33]. Slow cognitive processes run serially and require cognitive capacity to answer or address specific questions. In the context of eHealth literacy, the handling of health information from the internet clearly represents a serial process of subsequent cognitions that require different competencies building on each other.

eHEALS: Translations of the Original eHEALS Questionnaire and its Limitations

Since its publication, the original eHEALS questionnaire has been translated into many languages, including Italian [34,35], Spanish [36], Dutch [37], Chinese [38], Serbian [39], Korean [40], Indonesian [41] and German [42]. However, some of these studies could not confirm the 1-factorial model as assumed by Norman and Skinner [16]. Looking at many different validation studies of the eHEALS questionnaire, a consistent factorial structure has not been verified; 1-factorial [16,37,43], 2-factorial [42,44,45], and 3-factorial models [46-48] have been identified in different validation studies and languages. These results indicate that the eHEALS questionnaire lacks consistent factorial structure.

The German version of the questionnaire validated by Soellner and colleagues [42] especially lacks methodological and content-related accuracy. They developed an initial instrument for assessing eHealth literacy for the German-speaking community (G-eHEALS). However, Soellner and colleagues [42] did not meet scientific criteria substantially; first, they did not meet the criteria scientifically recommended for translation of instruments. Second, in their 2-factorial model content validity was questionable because some items reflected the subdimension of information appraisal rather than the assigned subdimension of information seeking ("I know how to use the health information I find on the internet to help me" or "I feel

confident in using information from the internet to make health decisions"). In addition, Soellner and colleagues [42] collected their data on a limited sample of 327 students aged 16 to 21 years at only one type of school (gymnasium: a German school type preparing for university attendance), and people of older age were not considered for validation. However, as people of older age may be less familiar using the internet [49-51] and eHealth literacy especially depicts a particular digital literacy, the model proposed by Soellner and colleagues [42] is possibly not valid for assessing eHealth literacy in older people. Moreover, the educational level of the participants could not be considered within their biased study sample. Juvalta and colleagues [52], who used the G-eHEALS, have also collected their data on a limited sample of young parents (88.5% female). In another German-speaking study, Reder and colleagues [53] have shown a 3-factorial structure for the G-eHEALS. However, only women participated in this study, which is a limited sample for examining the validity of the G-eHEALS. Inconsistent findings and methodological limitations of these studies indicate an unclear factorial structure of the G-eHEALS.

Another limitation of the original eHEALS questionnaire refers to insufficient representation of an elaborated cognitive information process. The original scale does not reflect the above-mentioned complexity of an information process in its entirety. Petrič and colleagues [54] focused on this limitation and developed an extended eHealth literacy scale (eHEALS-E). Creating a 20-item questionnaire, they found a 6-factorial structure. Despite this extension and other concepts and questionnaires [55-57], eHEALS is still the instrument most used for measuring eHealth literacy.

Aims of This Study

In summary, the G-eHEALS validated by Soellner and colleagues [42] was a valuable first approach to the important topic of eHealth literacy, but it underlies significant methodological limitations and lacks in psychometric quality. Nevertheless, as eHealth literacy could be confirmed as an important construct of health-related outcomes, the possibility of assessing eHealth literacy is crucial for health care practitioners and researchers in understanding health competence in German-speaking people. In response to the practical and scientific demands and described limitations, we developed a new instrument for measuring eHealth literacy with 4 objectives:

- Extension of the existing questionnaire of Norman and Skinner [16] by 8 nonoverlapping items proposed by Petrič and colleagues [54]. By combining the questionnaires, a better representation of the construct of eHealth literacy regarding the cognitive processes of seeking, identifying, and evaluating health information should be achieved.
- German translation of the items according to common scientific recommendations [58,59] to ensure content validity.
- Validation of the revised GR-eHEALS at a convenience sample in terms of construct and criterion validity. We decided to collect data in a convenience sample to reach participants with varied socioeconomic backgrounds. Furthermore, our goal was not to limit the sample in order to develop a measurement model that is as generic as possible.
- To our knowledge, there is no study examining measurement invariance of eHealth literacy between gender, age, or educational level in a German sample. Nevertheless, the interpretation of statistical differences between different groups of people requires measurement invariance between these groups [60]. As eHealth literacy represents competencies that are important for people regardless of their sociodemographic status, its measurement should obviously be independent of these influencing variables.

All in all, we are pursuing the study goals to develop a revised and validated instrument for measuring eHealth literacy. Further, we sought to examine the measurement invariance of the instrument regarding relevant sociodemographic variables.

Methods

Development of the New Instrument

The revised eHealth Literacy Scale (GR-eHEALS) is based on the original items from the eHEALS [16] extended by adding items from the eHEALS-E questionnaire from Petrič and colleagues [54]. The translation was conducted following the guidelines proposed by Beaton and colleagues [58] and Guillemin and colleagues [59] for translation of academic literature to ensure content validity. Accordingly, in a first step, 2 of the authors translated the items into German and merged these translations into a first translation proposal. In the second step, this proposal was discussed within a systematic expert panel consisting of the 2 translators and 2 psychologists who are experts in the context of health care and eHealth. The resulting second proposal was translated back into English in the third step to confirm that the essential meaning of the items is consistent with the original items. In the fourth step, cognitive interviews were conducted to make sure that all items are easy to understand, do not include offensive speech, and do not discriminate for age or gender. Interviewees were aged 23 to 72 years and had different educational backgrounds. The resulting final version of the translated and extended version consisted of 16 items. The original items and the translated items are displayed in Multimedia Appendix 1. Items 1 to 8 are translated from the original eHEALS questionnaire from Norman & Skinner [16], and items 9 to 16 are translated from the questionnaire (eHEALS-E) from Petrič and colleagues [54]. All subsequent nominations of item numbers refer to the item numbers mentioned in Multimedia Appendix 1. To validate the GR-eHEALS, we performed a prestudy in which we aimed to check for any complications in answering the translated items and to conduct an item analysis The results of this analysis are displayed in Multimedia Appendix 2. As the prestudy showed solid item characteristics, the developed instrument was considered good fitting for the purpose of the main study.

Study Design and Participants

The cross-sectional study was conducted via Unipark (Tivian XI GmbH), an online survey tool, between October and November 2020. The ethics committee of the Faculty of Medicine of the University of Duisburg–Essen reviewed and approved this study (20-9592-BO).

All data were collected anonymously. Participants for this study were recruited via personal and occupational networks and online social networks (Xing, Facebook, LinkedIn). In our analyses, only complete data sets were considered. From a total of 1634 participants, 524 have completed our questionnaire in full, which represents a completion rate of 32.1% and can be considered typical for an online survey [61]. We excluded cases in which participants took less than 5:34 minutes (5% percentile) or more than 25:45 minutes (95% percentile) to complete the survey. Furthermore, we excluded 1 participant for being under 18 years old. As only 1 person indicated gender as diverse, we excluded this case in order to perform the analysis of measurement invariance of gender. The resulting sample consisted of 470 respondents. The sample size is in accordance with recommendations for validation studies [62,63]. Answering the questionnaire took 11:32 (SD 4:24) minutes on average. All data supporting the conclusion of the study are included in Multimedia Appendix 3.

In the main study, it was our objective to validate the GR-eHEALS in a convenience sample to verify its convergent, discriminant, and criterion validity and test for measurement invariance.

We verified convergent validity by assuming a positive correlation between eHealth literacy and health literacy, which measures a similar construct but does not take the source of information into account. Furthermore, we assumed eHealth literacy to be positively interrelated with internet confidence and negatively associated with internet anxiety as eHealth literacy particularly focuses on the gathering of information from the internet.

To verify discriminant validity, we captured impulsivity and common personality traits assuming no significant interrelations. As eHealth literacy reflects competencies in dealing with health-related information [15] rather than a personality trait, there should be no content-related overlaps between eHealth literacy and personality traits.

Additionally, we considered the possible outcome variables mental and physical health status and life satisfaction to examine criterion validity. Criterion validity of an instrument describes the ability to prove relationships between the construct itself and possible outcomes [64]. Thus, we expected eHealth literacy to be associated with above mentioned health-related variables.

The survey included the following questionnaires (sample items presented below are translations). Most scales were assessed on 5-point Likert scales from 1=strongly disagree to 5=strongly agree. Exceptions are separately explained below. Scales contained inverted items that were recoded prior to statistical analyses.

Measurements

Health Literacy

Participants rated their health literacy on 16 items from the Health Literacy Questionnaire from Röthlin and colleagues [65]. A sample item is “How easy/difficult is it to find information about therapies for diseases that affect you?” Health literacy was measured on a 2-point scale (easy/hard). Therefore, it is used as a sum-score indicating the extent of health literacy between 0 and 16 (mean 12.63 [SD 2.99]). Cronbach alpha of this scale was .79.

Impulsivity

We used the 8-item Impulsive Behavior–8 Scale from Kovaleva and colleagues [66] to measure impulsivity (eg, “Sometimes I spontaneously do things that I should not have done”). Cronbach alpha of this scale was .72 (mean 2.78 [SD 0.59]).

Personality Traits

Personality traits (extraversion, neuroticism, openness, conscientiousness, and agreeableness) were each assessed by 2 items from Rammstedt and colleagues [67]. A sample item for neuroticism is “I get nervous and insecure easily.” Extraversion (mean 3.30 [SD 1.04]), neuroticism (mean 3.08 [SD 0.97]), openness (mean 3.61 [SD 0.99]), conscientiousness (mean 3.59 [SD 0.75]), and agreeableness (mean 3.15 [SD 0.76]) had Cronbach alphas of .79, .66, .62, .38, and .19, respectively. Due to low reliabilities, conscientiousness and agreeableness were excluded from the following analyses.

Further Constructs

In addition, we asked for internet confidence (3 items; mean 3.74 [SD 0.72], Cronbach alpha .89), internet anxiety (3 items; mean 1.81 [SD 0.82], Cronbach alpha .81) and single items to measure physical (mean 7.37 [SD 1.58]) and mental health (mean 7.27 [SD 1.90]) on 11-point Likert scales from 0=very bad health to 10=very good health (all self-formulated), and life satisfaction at a 5-point Likert scale from 1=not satisfied at all to 5=totally satisfied (mean 3.76 [SD 0.83]) from Beierlein and colleagues [68].

Furthermore, sociodemographic variables (age, gender, marital status, educational level, financial situation, internet availability, and community size) were considered to make sure that the sample represents the population.

Statistical Analysis

All data analyses were conducted using R (R Foundation for Statistical Computing), RStudio, and several packages.

Prior to conducting confirmatory factor analysis (CFA), we performed an exploratory factor analysis (EFA) to evaluate whether data were suitable for factor analysis. We used the Kaiser-Meyer-Olkin (KMO) and Bartlett test of sphericity for evaluation. Factor extraction was conducted using maximum likelihood estimation with Promax oblique rotation and number of factors were identified by scree plot inspection and Kaiser criterion (eigenvalue >1). Factor loadings ≥0.4 were considered as significant [69].

Subsequently, we performed consecutive CFA and compared fit indices and factor loadings to confirm the best-fitting model by considering the recommendations of Hu and Bentler [70] who assume to achieve a comparative fit index (CFI) and Tucker Lewis index (TLI) about 0.95 and root mean square error of approximation (RMSEA) and standardized root mean square residual (SRMR) about 0.06 and 0.08, respectively. We used the robust maximum likelihood estimator as our prestudy showed that items were slightly negative skewed, and a robust

estimator is more likely to produce less biased model statistics than maximum likelihood estimator [71].

Two-tailed Pearson correlations were conducted considering a significance level of 5% to examine convergent, discriminant, and criterion validity.

We performed tests of measurement invariance on our final model to examine whether the measurement is reliable for both genders as well as 2 age groups and 3 groups of educational level. For this purpose, we performed consecutive multigroup CFA with progressively stricter model assumptions by fixing an increasing number of model parameters for each of 3 measurement invariance models.

Measurement invariance—as a prerequisite for the interpretation of mean differences—is verified by 3 consecutive steps with increasingly strict model assumptions for (1) the number of factors and the pattern of factor-indicator relationships (configural invariance), (2) factor loadings (metric invariance), and (3) intercepts of indicators (scalar invariance) [72]. These 3 steps assume that there are no differences between observed groups regarding these parameters, and interpretation of mean differences is valid when scalar invariance is confirmed [73]. Differences between groups should only be interpreted when measurement invariance is confirmed since otherwise differences between groups may occur due to the fact that an instrument does not measure equally between different groups [60,73,74].

We applied a cutoff criterion of a difference of CFI (ΔCFI) of 0.01 as it is proposed as appropriate to assume invariance between two models [75,76]. Thus, for evaluation of measurement invariance we considered the model fit indices and difference of CFI between compared models.

Results

Sample Characteristics

Mean age of participants was 37.16 (SD 13.4, min 18, max 82, median 33) years. Sample characteristics of all other sociodemographic variables are shown in Table 1.

Table 1. Summary of sample characteristics (n=470).

Characteristics	Values, n (%)
Gender	
Female	332 (70.6)
Male	138 (29.4)
Marital status	
Married	161 (34.3)
Not married, in partnership	183 (38.9)
Single	115 (24.5)
Other	11 (2.3)
Educational level	
Lower secondary school	5 (1.1)
Upper secondary school	24 (5.1)
University entrance qualification	77 (16.4)
Vocational training	91 (19.4)
University degree	273 (58.1)
Financial situation	
Very good	9 (1.9)
Good	47 (10.0)
Middling	114 (24.3)
Bad	220 (46.8)
Very bad	80 (17.0)
Internet availability	
Always available	288 (61.3)
Mostly available	177 (37.7)
Occasionally available	5 (1.1)
Not available	0 (0.0)
Community size	
Big city (>100,000 inhabitants)	244 (51.9)
Medium city (>20,000 inhabitants)	88 (18.7)
Small city (>5000 inhabitants)	76 (16.2)
Rural village (<5000 inhabitants)	62 (13.2)

Exploratory Factor Analysis

KMO revealed a value of 0.92 and Bartlett test of sphericity was highly significant (P<.001), indicating that data were suitable for factor analysis. Empirical Kaiser criterion and scree plot implied a 2-factor model. Table 2 shows factor loadings of the 2 factors.

As item 14 did not significantly load on any of the 2 factors it was excluded from the following analysis. The remaining 15 items were considered in the CFA.

Table 2. Results of exploratory factor analysis.

Item no	Factor 1	Factor 2
1	0.88	–0.06
2	0.80	0.03
3	0.84	0.00
4	0.97	–0.06
5	0.49	0.37
6	0.10	0.62
7	0.03	0.70
8	0.28	0.49
9	0.00	0.78
10	–0.12	0.78
11	–0.11	0.75
12	–0.07	0.56
13	0.12	0.56
14	0.32	0.31
15	0.44	0.01
16	0.44	–0.09

Confirmatory Factor Analysis

In model 1, 15 items were assigned on the 2 factors identified by the EFA. Based on the content meanings of the underlying items, factor 1 represents information seeking and factor 2 represents information appraisal. However, items 13, 5, and 15 did not fit the factor proposed by the EFA in terms of their content. Therefore, item 13 was reassigned to information seeking whereas items 5 and 15 were reassigned to information appraisal in model 2. For model 3, we removed 6 items due to low factor loadings (<0.65). Moreover, we excluded 1 more item to develop a parsimonious model resulting in a 2-factorial model with 4 items on each of the 2 factors. Table 3 shows the model fits of the 3 models.

CFI, TLI, and SRMR practically meet the criteria of a good model fit. RMSEA is slightly above the recommendations of Hu and Bentler [70]. Considering the recommendations, model 3 shows an acceptable-to-good model fit.

Figure 1 depicts the structure of the 2-factorial model with its factor loadings. All item factor loadings were greater than λ=0.71.

Information seeking and information appraisal achieved satisfactory Cronbach alphas of .92 and .83, respectively. Table 4 shows the statistics of the final items. Based on mean and standard deviation, lower levels of information seeking and information appraisal are below a mean score of 2.99 and 3.20, respectively. Higher levels can be assumed above mean scores of 4.71 and 4.69, respectively.

Table 3. Results of the confirmatory factor analyses.

Model	Chi-square	df	CFI[a]	TLI[b]	RMSEA[c]	SRMR[d]	AIC[e]	BIC[f]
1	433.5	89	0.891	0.871	0.100	0.067	16029.832	16158.567
2	519.8	89	0.863	0.839	0.112	0.084	16136.608	16265.343
3	117.0	19	0.942	0.915	0.127	0.055	7782.043	7852.640

[a]CFI: comparative fit index.

[b]TLI: Tucker Lewis index.

[c]RMSEA: root mean square error of approximation.

[d]SRMR: standardized root mean square residual.

[e]AIC: Akaike information criterion.

[f]BIC: Bayesian information criterion.

Figure 1. A 2-factorial, intercorrelated model of eHealth literacy.

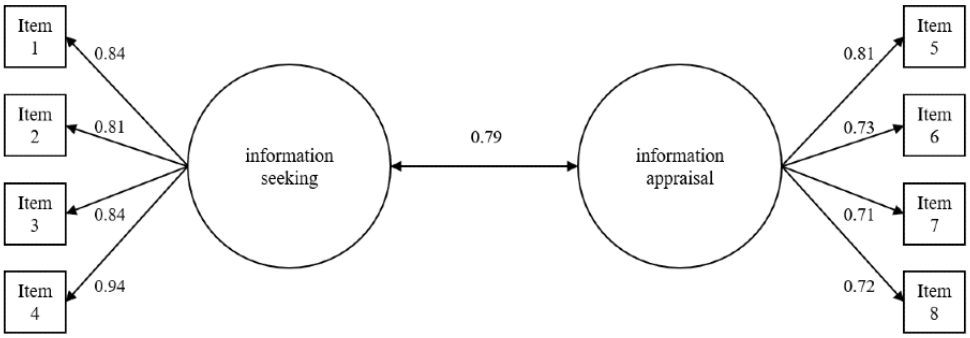


Table 4. Descriptive statistics of the revised German eHealth Literacy Scale (GR-eHEALS) items.

Item	Mean (SD)	Median	Skew
Information seeking	3.85 (0.86)	4.00	–0.78
1. Ich weiß, wie ich Internetseiten mit hilfreichen Gesundheitsinformationen finden kann.	3.93 (0.95)	4.00	–0.93
2. Ich weiß, wie ich das Internet nutzen kann, um Antworten auf meine Gesundheitsfragen zu erhalten.	4.04 (0.87)	4.00	–1.01
3. Ich weiß, welche Seiten mit Gesundheitsinformationen im Internet verfügbar sind.	3.63 (1.00)	4.00	–0.60
4. Ich weiß, wo ich im Internet hilfreiche Gesundheitsinformationen finden kann.	3.81 (1.01)	4.00	–0.89
Information appraisal	3.95 (0.74)	4.00	–0.77
5. Ich weiß Gesundheitsinformationen aus dem Internet so zu nutzen, dass sie mir weiterhelfen.	3.91 (0.88)	4.00	–0.77
6. Ich bin in der Lage, Internetseiten mit Gesundheitsinformationen kritisch zu bewerten.	4.18 (0.87)	4.00	–1.24
7. Ich kann zwischen vertrauenswürdigen und fragwürdigen Internetseiten mit Gesundheitsinformationen unterscheiden.	4.07 (0.84)	4.00	–0.93
8. Ich fühle mich sicher darin, Informationen aus dem Internet zu nutzen, um Entscheidungen in Bezug auf meine Gesundheit zu treffen.	3.62 (1.03)	4.00	–0.57

Validation of the GR-eHEALS

To examine convergent, discriminant, and criterion validity of the GR-eHEALS, we performed correlation analyses with the 2 factors (information seeking and information appraisal). Moreover, correlations of the 2 factors with demographic variables were calculated. Results are shown in Table 5. Both factors were strongly positively correlated with health literacy and internet confidence and strongly negatively correlated with internet anxiety. None of the 2 scales correlated significantly with impulsivity or extraversion. Information appraisal was interrelated with neuroticism while information seeking was associated with openness. Information appraisal was correlated with mental and physical health and life satisfaction, which was not true for information seeking. Furthermore, information seeking was significantly associated with age.

Table 5. Pearson correlation coefficients of the eHealth literacy factors.

Scales	Information seeking (*P* value)	Information appraisal (*P* value)
Convergent validity		
Health literacy	0.43 (<.001)	0.53 (<.001)
Internet confidence	0.17 (<.001)	0.17 (<.001)
Internet anxiety	–0.21 (<.001)	–0.23 (<.001)
Discriminant validity		
Impulsivity	–0.06 (.16)	–0.05 (.28)
Extraversion	–0.03 (.58)	0.03 (.56)
Neuroticism	–0.08 (.09)	–0.14 (.001)
Openness	0.10 (.03)	0.07 (.12)
Criterion validity		
Mental health	0.06 (.20)	0.19 (<.001)
Physical health	0.06 (.21)	0.12 (.01)
Life satisfaction	–0.01 (.83)	0.12 (.01)
Sociodemographic variables		
Age	0.10 (.02)	0.06 (.16)
Gender	–0.03 (.55)	0.01 (.78)
Marital status	–0.02 (.71)	–0.07 (.15)
Educational level	–0.04 (.39)	–0.02 (.68)
Financial situation	–0.05 (.27)	0.04 (.45)
Internet availability	0.01 (.76)	0.02 (.71)
Community size	0.02 (.60)	–0.04 (.41)

Test of Measurement Invariance

Measurement invariance of the GR-eHEALS was performed to test whether the scale is a suitable measurement independently of gender, age, and educational level. Prior to these analyses, a median split was performed to separate participants into 2 groups according to age. Median age was 33 years. Also, to divide the study sample into 3 groups of educational levels, we separated participants into people who held a university degree, people who completed a vocational training, and people who had any school certificate. Results of the analyses are shown in Table 6.

Besides chi-square and fit indices, Table 6 shows the differences of CFI between models. Regarding measurement invariance of gender and education, all changes in CFI are below 0.01, indicating that model fits did not substantially decrease between more constraint models. Measurement invariance regarding age must be rejected as configural invariance could not be confirmed.

Table 6. Results of measurement invariance for gender, age, and education using multigroup confirmatory factor analysis.

Model	Chi-square	df	CFI[a]	TLI[b]	RMSEA[c]	SRMR[d]	ΔCFI[e]
Gender[f]							
Configural[g]	154.937	38	0.94	0.905	0.135	0.056	0.006
Metric	166.889	44	0.93	0.916	0.128	0.066	0.002
Scalar	181.273	50	0.93	0.923	0.122	0.068	0.002
Age[h]							
Configural[g]	187.672	38	0.92	0.883	0.150	0.059	0.021
Metric	185.713	44	0.92	0.901	0.138	0.059	–0.002
Scalar	197.419	50	0.92	0.913	0.130	0.060	0.001
Education[i]							
Configural[g]	170.758	57	0.94	0.904	0.136	0.058	0.007
Metric	174.474	69	0.94	0.926	0.119	0.061	–0.004
Scalar	196.107	81	0.94	0.934	0.112	0.063	0.002

[a]CFI: comparative fit index.

[b]TLI: Tucker Lewis index.

[c]RMSEA: root mean square error of approximation.

[d]SRMR: standardized root mean square residual.

[e]Change in CFI compared to preceding model.

[f]Female n=332; male n=138.

[g]Change of CFI compared to model 3.

[h]Age>median n=240; age<median n=230.

[i]University degree n=273; vocational training n=91; school certificate n=106.

Discussion

Principal Findings

The results of our factor analyses show that eHealth literacy consists of 2 factors, information seeking and information appraisal. Our first study aim was to examine whether the measurement of eHealth literacy could be improved by adding nonoverlapping items from the eHEALS-E [54] to the original eHEALS [16]. We performed an EFA and several CFAs to examine the factorial structure of our instrument. Our analyses show that the measurement of eHealth literacy could not be improved by adding additional items to the well-established eHEALS questionnaire.

However, our study significantly contributes to the existing measurement of eHealth literacy. By strongly following scientific recommendations regarding academic translations, we developed the GR-eHEALS with high content validity. By taking statistical and content-related consideration into account when conducting factor analyses, we developed a measurement model of eHealth literacy with high content validity and acceptable-to-good model fit. Cronbach alpha was satisfactory for the 2 factors indicating good internal consistency and confirming reliability of the instrument.

Our findings on the examination of convergent, discriminant, and criterion validity of our instrument were not completely consistent with our expectations and require critical discussion. As expected, the 2 factors showed significant correlations with the convergent constructs of health literacy, internet confidence, and internet anxiety. By contrast, while impulsivity and extraversion consistently showed, as expected, no significant correlations with the 2 factors, neuroticism and openness indicated more inconsistent interrelations. Neuroticism was strongly negatively correlated with information appraisal, but not with information seeking. On the other hand, openness was only correlated with information seeking but not with information appraisal. To understand these unexpected correlational patterns, we examined findings of studies discovering the associations of personality traits and health-related constructs. Other studies showed that neuroticism is associated with lower health behavior self-efficacy and health behaviors [77] and lower internet use for learning and education [78]. These findings could indicate that neuroticism distorts cognitive processes of higher elaboration that are required for information appraisal but not necessarily for information seeking. Regarding the personality trait of openness, Bogg and Vo [79] have shown that people with higher openness more often search the internet regarding health-related topics. One could think that openness promotes people to search for new information in a sense of curiosity. However, the subsequent and cognitively demanding process of information appraisal may not be promoted by people's openness.

Referring to the examination of criterion validity, positive correlations with the possible outcome variable mental health,

physical health, and life satisfaction were expected, although only information appraisal was significantly related to these constructs. These results could be potentially explained by the idea that information seeking is a process that requires cognitive efforts but may not be sufficient to promote satisfaction and health status on its own but needs a high competency in information appraisal as a mandatory precondition. However, the search of information is a necessary process to perform the subsequent process of information appraisal.

To sum up, convergent validity of our instrument can be comprehensively confirmed. Examination of discriminant validity and criterion validity reveal unexpected findings that should be subjects of further studies. Despite our results not completely meeting our expectations, findings indicate that the 2 factors represent different cognitive processes in line with dual-process theories of analytic and rule-based processes: information seeking as a first of 2 consecutive competencies exclusively focuses on the process of searching information on the internet but not on a deeper application of the information found. Within a second consecutive competency built on information seeking, information appraisal describes a cognitive process of interpretation of information and its application on personal health-related questions.

Furthermore, we investigated the measurement invariance for gender, age, and educational level. The results of our study suggest that measurement invariance of the GR-eHEALS can be assumed for gender and educational level at a scalar level of invariance but not for age. Our study is the first to examine measurement invariance for these sociodemographic variables. Particularly regarding sample limitations of previous studies investigating eHealth literacy, the GR-eHEALS is the first instrument that can be deployed and interpreted regardless of gender and educational level. Therefore, future researchers are able to interpret statistical differences of these sociodemographic variables on eHealth literacy by using the GR-eHEALS. This is highly important as one could think of differential levels of eHealth literacy due to gender, which was confirmed for the construct of health literacy [80]. Regarding educational level, studies suggest that education also plays a role in the context of eHealth literacy [81,82], but, to our knowledge, neither used instruments confirmed to be measurement invariant.

Concerning the finding of inequality of our instrument with respect to age, one potential explanation could be that older people are less familiar with using the internet than younger people in terms of a digital divide [49] and have a different understanding of information seeking and information appraisal than younger people. Chesser and colleagues [83] suggest that age is a relevant variable in the context of eHealth literacy. Further, in our data we found significant interrelations of age and information seeking but not of age and information appraisal. This should be examined further in upcoming research.

In summary, prior research indicates that the investigation of differences of eHealth literacy of different groups of people is of high scientific interest. Nonetheless, previous studies were lacking considering statistical differences should not be interpreted unless measurement invariance is confirmed. With the GR-eHEALS, we close this gap and contribute substantially to the understanding of the concept of eHealth literacy and the interpretation of mean differences for gender and educational level.

Due to its high validity, the GR-eHEALS provides researchers and practitioners with a measurement for the increasingly important construct of eHealth literacy. As eHealth literacy is linked with many health-related outcomes and behaviors [19,26,27], the GR-eHEALS could provide a basis for educational programs to improve eHealth literacy by focusing on the main cognitive processes important for interpreting health information from the internet. Also, there is evidence that students lack in competencies regarding eHealth literacy [84]. Hence, the assessment and development of eHealth literacy should be a part of students' curriculum to provide young people with the competencies needed to maintain or improve one's health status. Consequently, the GR-eHEALS could be part of educational psychologists' diagnostic repertoire as well as a foundation for specialist training programs in schools and universities. We propose that the results of the GR-eHEALS should be interpreted based on the 2 competencies for diagnostic and interventions of eHealth literacy considering the described mean scores for higher and lower levels of information seeking and information appraisal.

Strengths and Limitations

The main strengths of this study are the high methodological and psychometric standards applied to develop GR-eHEALS and confirm its content, construct, and criterion validity. Furthermore, confirmation of measurement invariance is a state-of-the-art approach with strong practical implications regarding the interpretations of group differences.

One limitation of our study was that we measured eHealth literacy by self-assessment only. Since this construct is intended to measure skills and competencies, eHealth literacy should either be compared with actual behaviors or assessed using behavior-based measurement. Furthermore, our data were collected in a cross-sectional study. Therefore, correlational directions show relationships but are not interpretable causally. Future research should explore if our 2 factors show different effects on health-related outcomes. Additionally, as we used an online survey, participation by people familiar with the internet was more likely than by people who rarely use the internet. Thus, the possibility of selection bias should be considered. In our sample, a high proportion of people holding a university degree limits the representativeness regarding the education level. As in Germany about 19% of the population hold a university degree [85], our sample with a proportion of 58% holding a university degree clearly overrepresents academic persons. Even though it was our goal to collect data on a convenience sample, our study sample consisted of 71% female participants and cannot be considered as population-representative. Therefore, future studies should replicate our findings using a population-representative sample.

Conclusion

eHealth literacy reflects the important competence of people in maintaining and improving their health status. This competence

will become more and more important since the internet provides a rapidly increasing amount of health information with considerable bandwidth of quality and trustworthiness. The GR-eHEALS, with its 8 items on 2 factors, is a validated instrument to capture eHealth literacy in the German language. The GR-eHEALS contributes to the measurement of eHealth literacy in 3 ways: (1) instrument has high content validity because of a translation following scientific recommendations, (2) instrument has an acceptable-to-good model fit and confirms measurement invariance for gender and educational level, and (3) instrument revises the existing G-eHEALS and fills an important gap in measuring eHealth literacy to provide researchers and practitioners an accurate and valid assessment.

Acknowledgments

We acknowledge support by the Open Access Publication Fund of the University of Duisburg-Essen. The authors wish to thank Daniela Geiß and Maria Spank for their support as experts in the systematic expert panel within the translational process.

Authors' Contributions

MM, GE, AB, and EMS conceptualized the study. Project administration was performed by MM, GE, and AB. Statistical analyses were conducted by MM. MM and GE interpreted the data and wrote the original draft of the manuscript. AB, EMS, and MT supervised the project and contributed to the study design, data collection, and critical revision of the manuscript. All authors reviewed and approved the final manuscript. All data supporting the conclusion of the study are included in Multimedia Appendix 3.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Original and translated items.
[PDF File (Adobe PDF File), 105 KB - jmir_v24i2e28252_app1.pdf]

Multimedia Appendix 2
Item statistics of the prestudy (n=50).
[PDF File (Adobe PDF File), 102 KB - jmir_v24i2e28252_app2.pdf]

Multimedia Appendix 3
Dataset.
[XLSX File (Microsoft Excel File), 391 KB - jmir_v24i2e28252_app3.xlsx]

References

1. Parker RM, Baker DW, Williams MV, Nurss JR. The test of functional health literacy in adults: a new instrument for measuring patients' literacy skills. J Gen Intern Med 1995 Oct;10(10):537-541. [doi: 10.1007/BF02640361] [Medline: 8576769]
2. Nutbeam D. Health promotion glossary. Health Promot 1986 May;1(1):113-127. [doi: 10.1093/heapro/1.1.113] [Medline: 10318625]
3. Peterson PN, Shetterly SM, Clarke CL, Bekelman DB, Chan PS, Allen LA, et al. Health literacy and outcomes among patients with heart failure. JAMA 2011 Apr 27;305(16):1695-1701 [FREE Full text] [doi: 10.1001/jama.2011.512] [Medline: 21521851]
4. Schillinger D, Grumbach K, Piette J, Wang F, Osmond D, Daher C, et al. Association of health literacy with diabetes outcomes. JAMA 2002;288(4):475-482. [doi: 10.1001/jama.288.4.475] [Medline: 12132978]
5. World Health Organization. Shanghai declaration on promoting health in the 2030 Agenda for Sustainable Development. Health Promot Int 2017 Feb 01;32(1):7-8. [doi: 10.1093/heapro/daw103] [Medline: 28180270]
6. Cline RJ, Haynes KM. Consumer health information seeking on the Internet: the state of the art. Health Educ Res 2001 Dec;16(6):671-692. [doi: 10.1093/her/16.6.671] [Medline: 11780707]
7. Chen Y, Li C, Liang J, Tsai C. Health information obtained from the internet and changes in medical decision making: questionnaire development and cross-sectional survey. J Med Internet Res 2018 Dec 12;20(2):e47 [FREE Full text] [doi: 10.2196/jmir.9370] [Medline: 29434017]
8. Trotter MI, Morgan DW. Patients' use of the Internet for health related matters: a study of Internet usage in 2000 and 2006. Health Informatics J 2008 Sep;14(3):175-181. [doi: 10.1177/1081180X08092828] [Medline: 18775824]
9. Korp P. Health on the Internet: implications for health promotion. Health Educ Res 2006 Feb;21(1):78-86 [FREE Full text] [doi: 10.1093/her/cyh043] [Medline: 15994845]

10. Khaleel I, Wimmer BC, Peterson GM, Zaidi STR, Roehrer E, Cummings E, et al. Health information overload among health consumers: a scoping review. Patient Educ Couns 2020 Jan;103(1):15-32. [doi: 10.1016/j.pec.2019.08.008] [Medline: 31451363]
11. Ahmed ST. Managing news overload (MNO): the COVID-19 infodemic. Information 2020 Jul 25;11(8):375. [doi: 10.3390/info11080375]
12. Rathore FA, Farooq F. Information overload and infodemic in the COVID-19 pandemic. J Pak Med Assoc 2020 May;70(Suppl 3)(5):S162-S165. [doi: 10.5455/JPMA.38] [Medline: 32515403]
13. Lancet. The COVID-19 infodemic. Lancet Infect Dis 2020 Aug;20(8):875 [FREE Full text] [doi: 10.1016/S1473-3099(20)30565-X] [Medline: 32687807]
14. Orso D, Federici N, Copetti R, Vetrugno L, Bove T. Infodemic and the spread of fake news in the COVID-19-era. Eur J Emerg Med 2020 Apr 23:1 [FREE Full text] [doi: 10.1097/MEJ.0000000000000713] [Medline: 32332201]
15. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
16. Norman CD, Skinner HA. eHEALS: the eHealth literacy scale. J Med Internet Res 2006 Nov;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
17. Karnoe A, Kayser L. How is eHealth literacy measured and what do the measurements tell us? A systematic review. Knowl Manag E-Learn Int J 2015:576-600. [doi: 10.34105/j.kmel.2015.07.038]
18. Neter E, Brainin E. Association between health literacy, eHealth literacy, and health outcomes among patients with long-term conditions. Eur Psychol 2019 Jan;24(1):68-81. [doi: 10.1027/1016-9040/a000350]
19. Britt RK, Collins WB, Wilson KM, Linnemeier G, Englebert AM. The role of eHealth literacy and HPV vaccination among young adults: implications from a planned behavior approach. Comm Res Rep 2015 Jul 15;32(3):208-215. [doi: 10.1080/08824096.2015.1052963]
20. An L, Bacon E, Hawley S, Yang P, Russell D, Huffman S, et al. Relationship between coronavirus-related eHealth literacy and COVID-19 knowledge, attitudes, and practices among US adults: web-based survey study. J Med Internet Res 2021 Mar 29;23(3):e25042 [FREE Full text] [doi: 10.2196/25042] [Medline: 33626015]
21. Mitsutake S, Shibata A, Ishii K, Oka K. Association of eHealth literacy with colorectal cancer knowledge and screening practice among internet users in Japan. J Med Internet Res 2012;14(6):e153 [FREE Full text] [doi: 10.2196/jmir.1927] [Medline: 23149453]
22. Park H, Moon M, Baeg JH. Association of eHealth literacy with cancer information seeking and prior experience with cancer screening. Comput Inform Nurs 2014 Sep;32(9):458-463. [doi: 10.1097/CIN.0000000000000077] [Medline: 25105588]
23. Stellefson ML, Shuster JJ, Chaney BH, Paige SR, Alber JM, Chaney JD, et al. Web-based health information seeking and eHealth literacy among patients living with chronic obstructive pulmonary disease (COPD). Health Commun 2017 Sep 05:1-15. [doi: 10.1080/10410236.2017.1353868] [Medline: 28872905]
24. Lin C, Ganji M, Griffiths MD, Bravell ME, Broström A, Pakpour AH. Mediated effects of insomnia, psychological distress and medication adherence in the association of eHealth literacy and cardiac events among Iranian older patients with heart failure: a longitudinal study. Eur J Cardiovasc Nurs 2020 Feb;19(2):155-164. [doi: 10.1177/1474515119873648] [Medline: 31516036]
25. Noblin AM, Wan TTH, Fottler M. The impact of health literacy on a patient's decision to adopt a personal health record. Perspect Health Inf Manag 2012;9:1-13 [FREE Full text] [Medline: 23209454]
26. Mitsutake S, Shibata A, Ishii K, Oka K. Associations of eHealth literacy with health behavior among adult internet users. J Med Internet Res 2016 Jul;18(7):e192 [FREE Full text] [doi: 10.2196/jmir.5413] [Medline: 27432783]
27. Tsukahara S, Yamaguchi S, Igarashi F, Uruma R, Ikuina N, Iwakura K, et al. Association of eHealth literacy with lifestyle behaviors in university students: questionnaire-based cross-sectional study. J Med Internet Res 2020 Jun 24;22(6):e18155 [FREE Full text] [doi: 10.2196/18155] [Medline: 32579126]
28. Yang BX, Xia L, Huang R, Chen P, Luo D, Liu Q, et al. Relationship between eHealth literacy and psychological status during COVID-19 pandemic: a survey of Chinese residents. J Nurs Manag 2020 Dec 02:1 [FREE Full text] [doi: 10.1111/jonm.13221] [Medline: 33269492]
29. Li X, Liu Q. Social media use, eHealth literacy, disease knowledge, and preventive behaviors in the COVID-19 pandemic: cross-sectional study on chinese netizens. J Med Internet Res 2020 Oct 09;22(10):e19684 [FREE Full text] [doi: 10.2196/19684] [Medline: 33006940]
30. Evans JSBT, Stanovich KE. Dual-process theories of higher cognition: advancing the debate. Perspect Psychol Sci 2013 May;8(3):223-241. [doi: 10.1177/1745691612460685] [Medline: 26172965]
31. Kruglanski AW, Gigerenzer G. Intuitive and deliberate judgments are based on common principles. Psychol Rev 2011 Jan;118(1):97-109. [doi: 10.1037/a0020762] [Medline: 21244188]
32. Pennycook G, Rand DG. Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. J Pers 2020 Apr;88(2):185-200. [doi: 10.1111/jopy.12476] [Medline: 30929263]
33. Stanovich KE, West RF. Individual differences in reasoning: implications for the rationality debate? Behav Brain Sci 2000 Oct;23(5):645-665. [doi: 10.1017/s0140525x00003435] [Medline: 11301544]

34. De Caro W, Corvo E, Marucci AR, Mitello L, Lancia L, Sansoni J. eHealth literacy scale: an nursing analysis and Italian validation. Stud Health Technol Inform 2016;225:949. [Medline: 27332425]
35. Del Giudice P, Bravo G, Poletto M, De Odorico A, Conte A, Brunelli L, et al. Correlation between eHealth literacy and health literacy using the eHealth literacy scale and real-life experiences in the health sector as a proxy measure of functional health literacy: cross-sectional web-based survey. J Med Internet Res 2018 Oct 31;20(10):e281 [FREE Full text] [doi: 10.2196/jmir.9401] [Medline: 30381283]
36. Paramio PG, Almagro BJ, Hernando GA, Aguaded GJI. [Validation of the eHealth literacy scale (eHEALS) in Spanish university students]. Rev Esp Salud Publica 2015;89(3):329-338 [FREE Full text] [doi: 10.4321/S1135-57272015000300010] [Medline: 26388346]
37. van der Vaart R, van Deursen AJ, Drossaert CH, Taal E, van Dijk JA, van de Laar MA. Does the eHealth Literacy Scale (eHEALS) measure what it intends to measure? Validation of a Dutch version of the eHEALS in two adult populations. J Med Internet Res 2011;13(4):e86 [FREE Full text] [doi: 10.2196/jmir.1840] [Medline: 22071338]
38. Chang A, Schulz PJ. The measurements and an elaborated understanding of Chinese eHealth literacy (C-eHEALS) in chronic patients in China. Int J Environ Res Public Health 2018 Dec 23;15(7):1553. [doi: 10.3390/ijerph15071553] [Medline: 30041420]
39. Gazibara T, Cakic J, Cakic M, Pekmezovic T, Grgurevic A. eHealth and adolescents in Serbia: psychometric properties of eHeals questionnaire and contributing factors to better online health literacy. Health Promot Int 2018 May 25:770-778. [doi: 10.1093/heapro/day028] [Medline: 29800141]
40. Chung S, Park BK, Nahm E. The Korean eHealth Literacy Scale (K-eHEALS): reliability and validity testing in younger adults recruited online. J Med Internet Res 2018 Apr 20;20(4):e138 [FREE Full text] [doi: 10.2196/jmir.8759] [Medline: 29678800]
41. Wijaya MC, Kloping YP. Validity and reliability testing of the Indonesian version of the eHealth Literacy Scale during the COVID-19 pandemic. Health Informatics J 2021;27(1):1460458220975466 [FREE Full text] [doi: 10.1177/1460458220975466] [Medline: 33446030]
42. Soellner R, Huber S, Reder M. The concept of eHealth literacy and its measurement. J Media Psychol 2014 Jan;26(1):29-38. [doi: 10.1027/1864-1105/a000104]
43. Ma Z, Wu M. The psychometric properties of the Chinese eHealth Literacy Scale (C-eHEALS) in a Chinese rural population: cross-sectional validation study. J Med Internet Res 2019 Oct 22;21(10):e15720 [FREE Full text] [doi: 10.2196/15720] [Medline: 31642811]
44. Dale JG, Lüthi A, Fundingsland Skaraas B, Rundereim T, Dale B. Testing measurement properties of the Norwegian version of electronic Health Literacy Scale (eHEALS) in a group of day surgery patients. J Multidiscip Healthc 2020;13:241-247 [FREE Full text] [doi: 10.2147/JMDH.S242985] [Medline: 32210568]
45. Richtering SS, Morris R, Soh S, Barker A, Bampi F, Neubeck L, et al. Examination of an eHealth literacy scale and a health literacy scale in a population with moderate to high cardiovascular risk: Rasch analyses. PLoS One 2017;12(4):e0175372 [FREE Full text] [doi: 10.1371/journal.pone.0175372] [Medline: 28448497]
46. Hyde LL, Boyes AW, Evans T, Mackenzie LJ, Sanson-Fisher R. Three-factor structure of the eHealth Literacy Scale among magnetic resonance imaging and computed tomography outpatients: a confirmatory factor analysis. JMIR Hum Factors 2018 Feb 19;5(1):e6 [FREE Full text] [doi: 10.2196/humanfactors.9039] [Medline: 29459356]
47. Sudbury-Riley L, FitzPatrick M, Schulz PJ. Exploring the measurement properties of the eHealth Literacy Scale (eHEALS) among baby boomers: a multinational test of measurement invariance. J Med Internet Res 2017 Feb 27;19(2):e53 [FREE Full text] [doi: 10.2196/jmir.5998] [Medline: 28242590]
48. Stellefson M, Paige SR, Tennant B, Alber JM, Chaney BH, Chaney D, et al. Reliability and validity of the telephone-based eHealth Literacy Scale among older adults: cross-sectional survey. J Med Internet Res 2017 Oct 26;19(10):e362 [FREE Full text] [doi: 10.2196/jmir.8481] [Medline: 29074471]
49. Loges W, Jung J. Exploring the digital divide: internet connectedness and age. Commun Res 2016 Jun 30;28(4):536-562. [doi: 10.1177/009365001028004007]
50. Paige SR, Miller MD, Krieger JL, Stellefson M, Cheong J. Electronic health literacy across the lifespan: measurement invariance study. J Med Internet Res 2018 Jul 09;20(7):e10434 [FREE Full text] [doi: 10.2196/10434] [Medline: 29986848]
51. Tennant B, Stellefson M, Dodd V, Chaney B, Chaney D, Paige S, et al. eHealth literacy and Web 2.0 health information seeking behaviors among baby boomers and older adults. J Med Internet Res 2015;17(3):e70 [FREE Full text] [doi: 10.2196/jmir.3992] [Medline: 25783036]
52. Juvalta S, Kerry MJ, Jaks R, Baumann I, Dratva J. Electronic health literacy in Swiss-German parents: cross-sectional study of eHealth Literacy Scale unidimensionality. J Med Internet Res 2020 Mar 13;22(3):e14492 [FREE Full text] [doi: 10.2196/14492] [Medline: 32167476]
53. Reder M, Soellner R, Kolip P. Do women with high eHealth literacy profit more from a decision aid on mammography screening? Testing the moderation effect of the eHEALS in a randomized controlled trial. Front Public Health 2019;7:46 [FREE Full text] [doi: 10.3389/fpubh.2019.00046] [Medline: 30931291]
54. Petrič G, Atanasova S, Kamin T. Ill literates or illiterates? Investigating the eHealth literacy of users of online health communities. J Med Internet Res 2017 Oct 04;19(10):e331 [FREE Full text] [doi: 10.2196/jmir.7372] [Medline: 28978496]

55. Kayser L, Karnoe A, Furstrand D, Batterham R, Christensen KB, Elsworth G, et al. A multidimensional tool based on the eHealth literacy framework: development and initial validity testing of the eHealth Literacy Questionnaire (eHLQ). J Med Internet Res 2018 Feb 12;20(2):e36 [FREE Full text] [doi: 10.2196/jmir.8371] [Medline: 29434011]
56. Norgaard O, Furstrand D, Klokker L, Karnoe A, Batterham R, Kayser L, et al. The e-health literacy framework: a conceptual framework for characterizing e-health users and their interaction with e-health systems. Knowl Manag E-Learn Int J 2015;7(4):540. [doi: 10.34105/j.kmel.2015.07.035]
57. van der Vaart R, Drossaert C. Development of the digital health literacy instrument: measuring a broad spectrum of Health 1.0 and Health 2.0 skills. J Med Internet Res 2017 Jan 24;19(1):e27 [FREE Full text] [doi: 10.2196/jmir.6709] [Medline: 28119275]
58. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976) 2000 Dec 15;25(24):3186-3191. [doi: 10.1097/00007632-200012150-00014] [Medline: 11124735]
59. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993 Dec;46(12):1417-1432. [doi: 10.1016/0895-4356(93)90142-n] [Medline: 8263569]
60. Vandenberg RJ, Lance CE. A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. Org Res Meth 2016 Jun 29;3(1):4-70. [doi: 10.1177/109442810031002]
61. Nulty DD. The adequacy of response rates to online and paper surveys: what can be done? Assess Eval Higher Educ 2008 Jun;33(3):301-314. [doi: 10.1080/02602930701293231]
62. Rouquette A, Falissard B. Sample size requirements for the internal validation of psychiatric scales. Int J Methods Psychiatr Res 2011 Dec;20(4):235-249 [FREE Full text] [doi: 10.1002/mpr.352] [Medline: 22020761]
63. Kyriazos TA. Applied psychometrics: sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. Psychology 2018;09(08):2207-2230. [doi: 10.4236/psych.2018.98126]
64. Cronbach L, Meehl P. Construct validity in psychological tests. Psychol Bull 1955 Jul;52(4):281-302. [doi: 10.1037/h0040957] [Medline: 13245896]
65. Röthlin F, Pelikan J, Ganahl K. Die Gesundheitskompetenzder 15-jährigen Jugendlichen in Österreich. Abschlussbericht der österreichischen Gesundheitskompetenz Jugendstudie im Auftrag des Hauptverbandsder österreichischen Sozialversicherungsträger (HVSV). Wien: Ludwig Boltzmann Gesellschaft GmbH; 2013. URL: https://www.sozialversicherung.at/cdscontent/load?contentid=10008.715507&version [accessed 2022-01-14]
66. Kovaleva A, Beierlein C, Kemper C, Rammstedt B. Die Skala Impulsives-Verhalten-8 (I-8). Zusammenstellung Sozialwissenschaftlicher Items Skalen ZIS 2014:1-22 [FREE Full text] [doi: 10.6102/ZIS183]
67. Rammstedt B, Kemper C, Klein M, Beierlein C, Kovaleva A. Big Five Inventory (BFI-10). Zusammenstellung Sozialwissenschaftlicher Items Skalen ZIS 2014:1-23 [FREE Full text] [doi: 10.6102/ZIS76]
68. Beierlein C, Kovaleva A, László Z, Kemper C, Rammstedt B. Kurzskala zur Erfassung der Allgemeinen Lebenszufriedenheit (L-1). Zusammenstellung Sozialwissenschaftlicher Items Skalen ZIS 2015:1-18 [FREE Full text] [doi: 10.6102/ZIS229]
69. Field A. Discovering Statistics Using IBM SPSS Statistics. 4th Edition. Los Angeles: SAGE Publications Ltd; 2013.
70. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Structural Eq Modeling Multidisc J 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
71. Finney S, DiStefano C. Nonnormal and categorical data in structural equation modeling. In: Hancock GR, Mueller RO, editors. Structural Equation Modeling, Second Course. 2nd Edition. Charlotte: IAP Information Age Publishing; 2013:439-492.
72. Widaman K, Reise S. Exploring the measurement invariance of psychological instruments: applications in the substance use domain. In: Bryant KJ, Windle M, West SG, editors. The Science of Prevention: Methodological Advances From Alcohol and Substance Abuse Research. Washington: American Psychological Association; 1997:281-324.
73. Dimitrov DM. Testing for factorial invariance in the context of construct validation. Meas Eval Couns Devel 2017 Mar 13;43(2):121-149. [doi: 10.1177/0748175610373459]
74. Milfont T, Fischer R. Testing measurement invariance across groups: applications in cross-cultural research. Int J Psychol Res 2010 Jun 30;3(1):111-130. [doi: 10.21500/20112084.857]
75. Hirschfeld G, Von Brachel R. Improving multiple-group confirmatory factor analysis in R? A tutorial in measurement invariance with continuous and ordinal indicators. Pract Assess Res Eval 2014;19(1):7. [doi: 10.1017/9781108750561.006]
76. Putnick DL, Bornstein MH. Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. Dev Rev 2016 Sep;41:71-90 [FREE Full text] [doi: 10.1016/j.dr.2016.06.004] [Medline: 27942093]
77. Williams PG, O'Brien CD, Colder CR. The effects of neuroticism and extraversion on self-assessed health and health-relevant cognition. Personality Indiv Diff 2004 Jul;37(1):83-94. [doi: 10.1016/j.paid.2003.08.001]
78. Tuten TL, Bosnjak M. Understanding differences in web usage: the role of need for cognition and the five factor model of personality. Soc Behav Pers 2001 Jan 01;29(4):391-398. [doi: 10.2224/sbp.2001.29.4.391]

79. Bogg T, Vo PT. Openness, neuroticism, conscientiousness, and family health and aging concerns interact in the prediction of health-related Internet searches in a representative U.S. sample. Front Psychol 2014;5:370 [FREE Full text] [doi: 10.3389/fpsyg.2014.00370] [Medline: 24808880]
80. Lee HY, Lee J, Kim NK. Gender differences in health literacy among Korean adults: do women have a higher level of health literacy than men? Am J Mens Health 2015 Sep;9(5):370-379. [doi: 10.1177/1557988314545485] [Medline: 25122719]
81. Xesfingi S, Vozikis A. eHealth literacy: in the quest of the contributing factors. Interact J Med Res 2016 May 25;5(2):e16 [FREE Full text] [doi: 10.2196/ijmr.4749] [Medline: 27226146]
82. Levin-Zamir D, Lemish D, Gofin R. Media Health Literacy (MHL): development and measurement of the concept among adolescents. Health Educ Res 2011 Apr;26(2):323-335. [doi: 10.1093/her/cyr007] [Medline: 21422003]
83. Chesser A, Burke A, Reyes J, Rohrberg T. Navigating the digital divide: a systematic review of eHealth literacy in underserved populations in the United States. Inform Health Soc Care 2016;41(1):1-19. [doi: 10.3109/17538157.2014.948171] [Medline: 25710808]
84. Stellefson M, Hanik B, Chaney B, Chaney D, Tennant B, Chavarria EA. eHealth literacy among college students: a systematic review with implications for eHealth education. J Med Internet Res 2011 Dec;13(4):e102 [FREE Full text] [doi: 10.2196/jmir.1703] [Medline: 22155629]
85. Autorengruppe Bildungsberichterstattung. Bildung in Deutschland 2020: Ein indikatorengestützter Bericht mit einer Analyse zu Bildung in einer digitalisierten Welt. Bielefeld: wbv Media GmbH & Co KG; 2020.

Abbreviations

CFA: confirmatory factor analysis
CFI: comparative fit index
EFA: exploratory factor analysis
eHEALS: eHealth Literacy Scale
eHEALS-E: extended eHealth Literacy Scale
G-eHEALS: German eHealth Literacy Scale
GR-eHEALS: revised German eHealth Literacy Scale
KMO: Kaiser-Meyer-Olkin test
RMSEA: root mean square error of approximation
SRMR: standardized root mean square residual
TLI: Tucker Lewis index

Edited by G Eysenbach; submitted 26.02.21; peer-reviewed by L Sudbury-Riley, E Roehrer, JR Bautista; comments to author 08.06.21; revised version received 24.07.21; accepted 19.11.21; published 02.02.22.

Please cite as:
Marsall M, Engelmann G, Skoda EM, Teufel M, Bäuerle A
Measuring Electronic Health Literacy: Development, Validation, and Test of Measurement Invariance of a Revised German Version of the eHealth Literacy Scale
J Med Internet Res 2022;24(2):e28252
URL: https://www.jmir.org/2022/2/e28252
doi:10.2196/28252
PMID:35107437

Original Paper

Novel Interactive Tool for Breast and Ovarian Cancer Risk Assessment (Bright Pink Assess Your Risk): Development and Usability Study

Elizabeth A Hibler[1], MPH, PhD; Angela J Fought[1], MS; Kiarri N Kershaw[1], PhD; Rebecca Molsberry[1], MPH; Virginia Nowakowski[1], MPH; Deborah Lindner[2], MD

[1]Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[2]Bright Pink, Chicago, IL, United States

Corresponding Author:
Elizabeth A Hibler, MPH, PhD
Department of Preventive Medicine
Feinberg School of Medicine
Northwestern University
680 N Lake Shore Drive, Suite 1400
Chicago, IL, 60611
United States
Phone: 1 3125031178
Email: elizabeth.hibler@northwestern.edu

Abstract

Background: The lifetime risk of breast and ovarian cancer is significantly higher among women with genetic susceptibility or a strong family history. However, current risk assessment tools and clinical practices may identify only 10% of asymptomatic carriers of susceptibility genes. Bright Pink developed the Assess Your Risk (AYR) tool to estimate breast and ovarian cancer risk through a user-friendly, informative web-based quiz for risk assessment at the population level.

Objective: This study aims to present the AYR tool, describe AYR users, and present evidence that AYR works as expected by comparing classification using the AYR tool with gold standard genetic testing guidelines.

Methods: The AYR is a recently developed population-level risk assessment tool that includes 26 questions based on the National Comprehensive Cancer Network (NCCN) guidelines and factors from other commonly used risk assessment tools. We included all women who completed the AYR between November 2018 and January 2019, with the exception of self-reported cancer or no knowledge of family history. We compared AYR classifications with those that were independently created using NCCN criteria using measures of validity and the McNemar test.

Results: There were 143,657 AYR completions, and most participants were either at increased or average risk for breast cancer or ovarian cancer (137,315/143,657, 95.59%). Using our estimates of *increased* and *average* risk as the gold standard, based on the NCCN guidelines, we estimated the sensitivity and specificity for the AYR algorithm–generated risk categories as 100% and 89.9%, respectively (P<.001). The specificity improved when we considered the additional questions asked by the AYR to define *increased* risk, which were not examined by the NCCN criteria. By race, ethnicity, and age group; we found that the lowest observed specificity was for the Asian race (85.9%) and the 30 to 39 years age group (87.6%) for the AYR-generated categories compared with the NCCN criteria.

Conclusions: These results demonstrate that Bright Pink's AYR is an accurate tool for use by the general population to identify women at increased risk of breast and ovarian cancer. We plan to validate the tool longitudinally in future studies, including the impact of race, ethnicity, and age on breast and ovarian cancer risk assessment.

(J Med Internet Res 2022;24(2):e29124) doi:10.2196/29124

KEYWORDS

breast cancer; ovarian cancer; risk assessment; genetic testing

Introduction

In the United States, there are nearly 250,000 cases of breast cancer and 25,000 cases of ovarian cancer diagnosed annually [1,2]. Breast cancer is the leading cause of cancer-related deaths and the most commonly occurring cancer among women globally [3-5]. In contrast, ovarian cancer is only the 10th most common cancer but the fifth leading cause of cancer-related death among women [1,6] and the leading cause of death from gynecological cancer [7]. Studies have estimated that 5% to 10% of breast cancer and 10% to 18% of ovarian cancer are due to hereditary susceptibility, in particular from breast cancer gene (*BRCA*) mutations or a strong family history [8-10]. Among *BRCA1* carriers, specifically, estimates suggest a 40% to 87% cumulative breast cancer risk by age 70 years, whereas for ovarian cancer risk, estimates range from 16% to 68% [11]. However, studies have also demonstrated that current risk assessment and clinical practices only identify a small proportion of the at-risk population.

A central component of breast and ovarian cancer prevention programs is education on personal risk, as determined by factors such as family history, genetic susceptibility, hormonal risk factors, and modifiable health behaviors [12-14]. However, studies have estimated that current practices identify at most 10% of asymptomatic BRCA 1/2 carriers [8,10,15-17]. Moreover, among those with a strong family history of breast or ovarian cancer, but who have not received genetic testing, an estimated 20% to 30% have a pathogenic mutation in a breast cancer susceptibility gene [10]. Bright Pink is a nonprofit organization devoted to improving breast and ovarian cancer prevention through the belief that empowering educated patients leads to (1) lower distress associated with cancer risk assessment and (2) improved cancer prevention outcomes [18,19].

To achieve these goals, Bright Pink developed the Assess Your Risk (AYR) tool that provides a user-friendly, web-based survey to determine lifetime breast and ovarian cancer risk at the population level. Bright Pink was developed in collaboration with genetic counselors, public health practitioners, and clinicians to ensure alignment with standards of breast and ovarian cancer risk assessment but with a modern, consumer-facing interface that would be accessible to the general population. Since 2014, the Bright Pink AYR has been used by more than 1.2 million women to assess breast and ovarian cancer risk. The AYR tool offers a streamlined, educational, and actionable user experience with the goal of substantially increasing the number of women nationwide, as well as internationally, learning about breast and ovarian cancer risk.

However, the AYR tool has not yet been compared with the gold standard clinical guidelines for identifying women at increased risk for breast and ovarian cancer who are eligible for consultation with a genetic counselor. The aims of this study are to (1) describe the AYR tool and its content, (2) describe the AYR tool user population, and (3) compare the AYR tool classification against the National Comprehensive Cancer Network (NCCN) guidelines for genetic and family high-risk assessment in breast and ovarian cancer [2,17]. Moreover, improving health equity and early identification of young adult women at increased risk are the primary goals of Bright Pink and public health. Thus, this analysis is also the first to examine AYR classification compared with the NCCN criteria by race, ethnicity, and age group. Demonstration that Bright Pink AYR is an accurate tool for risk assessment at the population level is a critical step in ensuring that the general population, as well as patients and providers, have a reliable tool for the identification and education of women at risk for breast and ovarian cancer.

Methods

Study Population

Bright Pink, a national nonprofit organization based in Chicago, Illinois, was started in 2007 with the mission of empowering women to know their breast and ovarian cancer risk to manage their health proactively. For this study, eligible participants included Bright Pink AYR 3.0 users from October 2018 to March 2019, who were women aged ≥18 years. We included women of White non-Hispanic or Latinx, African American, Asian, or Hispanic or Latina race and ethnicity and all age groups. We excluded participants with a personal history of breast or ovarian cancer (excluding nonmelanoma skin cancer) or no knowledge of family history from calculations of agreement between the NCCN and the AYR. The goal of the NCCN criteria is to identify individuals who may be eligible for genetic testing and counseling for breast or ovarian cancer; therefore, we included women at increased risk for either breast or ovarian cancer.

Ethics Approval

This study was approved by the Northwestern University Human Subjects Protection Program (review number STU00207333).

Bright Pink AYR: History

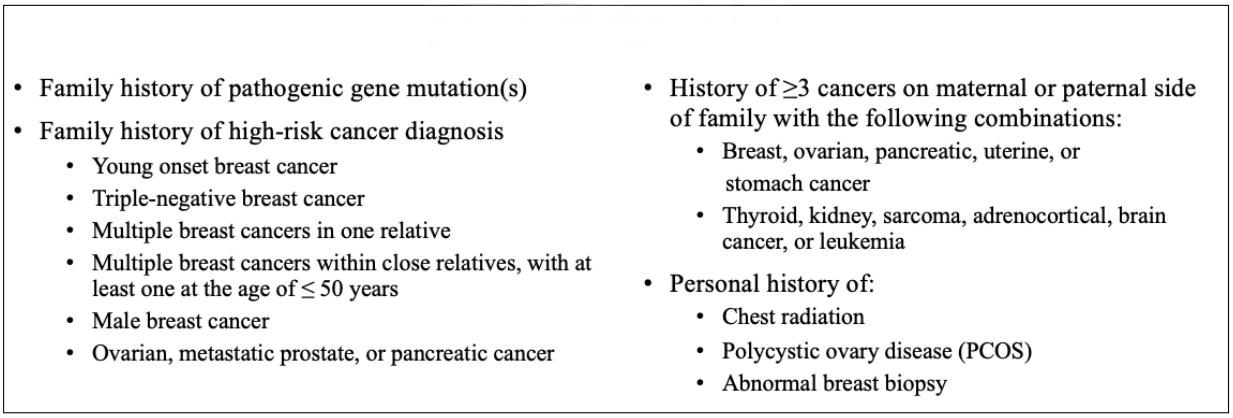
Bright Pink developed the original AYR tool in 2014 to support the goals of modernizing and increasing access to risk assessment for breast and ovarian cancer at the population level. The AYR tool estimates breast and ovarian cancer risk through a user-friendly, informative web-based quiz with superior-rated readability [20]. In late 2018, Bright Pink introduced a new version of AYR with an updated user experience and the goal of significantly increasing the number of women in the general population learning about breast and ovarian cancer risk. Specifically, the new version of the AYR tool includes a user-friendly quiz experience and updated mobile-first design features based on feedback gathered from surveys and focus groups of key target demographics, including African American women. Bright Pink also updated the educational content throughout the AYR tool to improve the user's understanding of how their responses affect risk calculations and understand personalized risk management recommendations. The National Society of Genetic Counselors has reviewed and approved the AYR tool. The AYR tool not only provides accessible risk assessment to the general population but also links women to additional resources, including Bright Pink's Explore Your Genetics website, breast health mobile messaging program, and an online peer support forum.

Bright Pink AYR: Design and Content

The AYR 3.0 tool includes 26 questions (Figure 1) based on the NCCN criteria for *Genetic/Familial High-Risk Assessment: Breast and Ovarian* [2,17]. However, the AYR tool is unique compared with other existing risk assessment tools designed primarily for use in clinical settings. The NCCN criteria that trigger consideration of genetic testing among asymptomatic individuals include (1) a family history of *BRCA1/2* or other gene variants, (2) a family history of high-risk cancers such as triple-negative or male breast cancer, or (3) a family history of more than 3 cases of cancers on either side.

On the basis of these NCCN criteria, Bright Pink's AYR tool categorizes women into the *high* category when they report already having a mutation in a breast cancer susceptibility gene identified through genetic testing. Specifically, AYR high risk is triggered by a personal history of positive genetic testing for genes associated with hereditary breast and ovarian cancer syndrome (*BRCA1* and *BRCA2*), Lynch syndrome (*MSH2*, *MLH1*, or *EPCAM*), or Peutz-Jeghers Syndrome (*STK11*) [2]. In contrast, *increased* risk is triggered by a personal history of positive genetic testing for other moderate to lower penetrance mutations (eg, *BRIP1*, *RAD51C*, *RAD51D*, *PMS2*, *and MSH6*) or a relative with a positive gene mutation test combined with no personal history of genetic testing. Moreover, a strong family history of high-risk breast or ovarian cancer, as well as a history of three or more cancers within the family, will trigger an increased risk in the AYR tool. Finally, AYR also uses components from the Gail and Tyrer-Cusick models, including a personal history of childhood radiation to the chest, abnormal breast biopsy, or polycystic ovary syndrome (PCOS) [21-24]. Women with no family history of these genes, and none of the other criteria listed in Figure 1, are categorized as *average* risk and provided with educational messaging focused on appropriate clinical screening guidelines and current evidence related to the impact of health behaviors on cancer prevention.

Figure 1. Description of the 3.0 criteria used by the Bright Pink Assess Your Risk (AYR) for breast and ovarian cancer risk assessment among asymptomatic women.



Statistical Methods

The primary goal of the AYR is to estimate the lifetime risk of breast and ovarian cancer among women with no history of breast or ovarian cancer. We thus excluded women reporting a previous history of breast or ovarian cancer from the analysis because the focus of AYR is on cancer prevention and early detection. In this subset, we calculated descriptive statistics for the AYR participants overall then by average, increased, and high AYR breast and ovarian categories. We then examined measures of agreement between the NCCN and AYR criteria.

We used the AYR quiz questions to create our own *average* and *increased* categories based on the NCCN criteria for breast and ovarian cancer genetic testing. We then compared these with the categories assigned by the AYR algorithm when the quiz was taken. We assessed measures including sensitivity and specificity to estimate the ability of AYR to distinguish between women who are eligible for genetic testing (AYR increased risk) versus those who are at average risk and thus not eligible for genetic testing, considering our NCCN criteria variables as the *gold standard*. We also examined the positive predictive value (PPV) and negative predictive value (NPV) to determine whether the AYR tool correctly identifies eligible or ineligible women based on NCCN criteria among those in the AYR increased or average categories, respectively. A *true positive* was when AYR risk category matched with the independently created NCCN categories. We calculated McNemar test of agreement, with an α of .05.

Results

AYR User Population

There were 143,657 AYR completions between October 2018 and March 2019, with complete data on health behaviors and family history. We examined the results for both breast (Table 1) and ovarian (Multimedia Appendix Table S1) cancer AYR risk categories. Overall, most AYR participants reported White race, and the mean age was 30.0 years (SD 10.7), with over half of participants in the 18-29 years age group. The mean BMI was 28.2 (SD 7.6) kg/m^2, and most women reported that they limit alcohol intake to less than 2 drinks per day and did not smoke. However, 60.54% (86,969/143,657) of women reported that they did not meet the physical activity guidelines for 150 minutes of moderate-intensity activity per week.

Table 1. Characteristics of the Bright Pink Assess Your Risk population for breast cancer risk.

Characteristics	Overall (N=143,657)	Average (n=62,760)	Increased (n=74,555)	High (n=6342)
Age (years), mean (SD)	29.66 (10.71)	29.8 (10.76)	30.2 (10.49)	29.0 (12.46)
Age group (years), n (%)				
18-29	84,506 (58.82)	37,689 (60.05)	42,533 (57.05)	4284 (67.55)
30-39	32,408 (22.56)	13,558 (21.60)	17,937 (24.06)	913 (14.40)
40-49	17,450 (12.15)	7369 (11.74)	9529 (12.78)	552 (8.70)
50-64	8152 (5.67)	3647 (5.86)	4050 (5.43)	455 (7.17)
≥65	1141 (0.79)	497 (0.79)	506 (0.68)	138 (2.18)
Race and ethnicity, n (%)				
White	97,227 (67.68)	38,984 (62.12)	53,694 (72.02)	4549 (71.73)
African American	5001 (3.48)	2620 (4.17)	2224 (2.98)	157 (2.48)
Asian	4469 (3.11)	2924 (4.66)	1398 (1.88)	147 (2.32)
Other or multiple	22,411 (15.60)	10,238 (16.31)	11,200 (15.02)	973 (15.34)
Ashkenazi Jewish ethnicity	716 (0.50)	249 (0.40)	400 (0.54)	67 (1.06)
Hispanic or Latina ethnicity	13,833 (9.62)	7745 (12.34)	5639 (7.56)	449 (7.08)
BMI, mean (SD)	28.22 (7.63)	28.00 (7.52)	28.51 (7.75)	27.12 (6.99)
Alcohol intake; ≥2 drinks per day, n (%)	16,171 (11.26)	6820 (10.87)	8498 (11.4)0	853 (13.45)
Exercise; <150 min/week, n (%)	86,969 (60.54)	37,481 (59.72)	45,879 (61.45)	3609 (56.91)
Current smoker, n (%)	21,801 (15.18)	8618 (13.73)	12,102 (16.23)	1081 (17.05)
Personal history of breast cancer only, n (%)	4599 (3.20)	0 (0)	0 (0)	4599 (72.52)
Personal history of ovarian cancer only, n (%)	1275 (0.89)	316 (0.50)	948 (1.27)	11 (0.17)
Personal history of breast and ovarian cancer, n (%)	1182 (0.82)	0 (0)	0 (0)	1182 (18.64)
Dense breasts, n (%)	24,347 (16.95)	8448 (13.46)	14,321 (19.21)	1578 (24.88)
History of breastfeeding, n (%)	24,720 (41.50)	10,554 (41.93)	13,304 (41.07)	862 (42.86)
Polycystic ovary syndrome, n (%)	13,825 (9.62)	5116 (8.15)	7881 (10.57)	828 (13.06)
Abnormal biopsy, n (%)	3521 (2.45)	0 (0)	2550 (3.42)	971 (15.31)
Chest radiation, n (%)	1541 (1.07)	0 (0)	1283 (1.72)	258 (4.07)
Family history early-onset breast cancer, n (%)	26,080 (18.15)	0 (0)	23,450 (31.45)	2630 (41.47)
Family history triple-negative breast cancer, n (%)	5851 (4.07)	0 (0)	5267 (7.06)	584 (9.21)
Family history multiple breast cancers in same relative, n (%)	16,453 (11.45)	0 (0)	14,888 (19.97)	1565 (24.68)
Family history of multiple breast cancers with at least one ≤50 years, n (%)	13,721 (9.55)	0 (0)	12,322 (16.53)	1399 (22.06)
Family history of male breast cancer, n (%)	1002 (0.70)	0 (0)	874 (1.17)	128 (2.02)
Family history of ovarian cancer, n (%)	16,711 (11.63)	0 (0)	15,704 (21.06)	1007 (15.88)
Family history of metastatic prostate cancer, n (%)	7813 (5.44)	0 (0)	7481 (10.03)	332 (5.23)
Family history of pancreatic cancer, n (%)	11,732 (8.17)	0 (0)	11,158 (14.97)	574 (9.05)
Personal genetic testing history, n (%)	5290 (3.68)	845 (1.35)	3064 (4.11)	1381 (21.78)
Family member genetic testing history, n (%)	20,553 (24.13)	2103 (6.46)	16,690 (34.26)	1750 (45.44)

We also examined participant characteristics by the risk category assigned by the AYR algorithm, which is shown for breast (Table 1) and ovarian (Multimedia Appendix 1, Table S1) cancers. Most participants were classified as average or increased risk for breast cancer, with only 4.41% (6342/143,657) reporting high risk. White race was the most commonly reported (97,227/143,657, 67.67%) compared with 3.48% (5001/143,657) for African American, 3.11% (4469/143,657) for Asian, and 9.62% (13,833/143,657) for Hispanic or Latina ethnicity. After stratifying for the AYR-assigned risk category, a greater

proportion of high-risk women tended to be aged >50 years and reported Ashkenazi Jewish ethnicity. For health behaviors, we observed a trend of higher alcohol intake and current smoking among women with increased and high breast cancer risk AYR categories. However, women in the high-risk category were least likely to report that they did not meet the physical activity guidelines (378/683, 55.3%) compared with 61.89% (49,566/80,091) and 58.53% (36,358/62,122) for increased and average risk, respectively) and reported the lowest mean BMI of 27.1 kg/m^2. High-risk women were also more likely to have a personal history of dense breast, PCOS, or abnormal breast biopsy as well as a family history of high-risk cancers. For ovarian cancer (Multimedia Appendix 1, Table S1), we observed similar trends overall and by risk category. This is supported by Table S2 in Multimedia Appendix 1, which demonstrates that most participants matched in the AYR risk category for breast and ovarian cancer.

Agreement Between AYR-Defined Categories Versus Independently Assessed NCCN Risk Categories

In examining the agreement between the AYR tool classifications and our classification based on NCCN criteria for breast and ovarian cancer risk, we calculated the sensitivity, specificity, PPV, and NPV. As shown in Table 2, the sensitivity and NPV were 100% for AYR compared with NCCN for all categories. However, overall, the specificity of the AYR tool was 89.9%, whereas the PPV was 94%. We also examined variation in measures of validity between the AYR and NCCN criteria by race, ethnicity, and age group (Table 3). We found the lowest estimates of specificity and PPV for Asian race, followed by Ashkenazi Jewish ancestry. Overall, specificity ranged from 85.9% for the Asian race to 91.3% for the African American race. For age group, there was also some variation in specificity, and the values ranged from 87.6% for the 30-39 years age group to 93.7% for those aged ≥65 years.

To explore the reasons for the observed variation in specificity and sensitivity, we also examined the frequency of the additional characteristics evaluated by the AYR tool, including personal history of chest radiation, PCOS, or abnormal breast biopsy by race, ethnicity, and age group (Multimedia Appendix 1, Table S3). Women with Ashkenazi Jewish ethnicity reported the highest frequency of previous chest radiation, PCOS, and abnormal breast biopsy. However, Asian women and women reporting for other races also reported 14.88% (638/4288) and 13.93% (2434/17,474) relatively high frequency, respectively, of a history of PCOS. Older women were more likely to report a history of chest radiation, whereas younger women were more likely to report PCOS. We also observed a trend of higher frequency of abnormal breast biopsy for the over 40-49, 50-64, and ≥65 years age groups (905/18,671, 4.85%; 644/7997, 8.05%; and 109/1151, 9.47%, respectively) compared with younger age groups.

Table 2. Variation in agreement between National Comprehensive Cancer Network (NCCN) and Assess Your Risk (AYR) by race, ethnicity, and age group.

	AYR compared with NCCN criteria only							
	Total[a], n (%)	True positive, n	True negative, n	Sensitivity (%)	Specificity, (%)	PPV[b] (%)	NPV[c] (%)	*P* value[d]
Overall	117,595 (100)	72,072	40,880	100.00	89.80	93.95	100.00	<.001
Race and ethnicity								
White	82,128 (69.83)	52,078	26,994	100.00	89.83	94.46	100.00	<.001
African American	3886 (3.3)	2125	1608	100.00	91.31	93.28	100.00	<.001
Asian	3187 (2.71)	1311	1612	100.00	85.93	83.24	100.00	<.001
Other	17,614 (14.97)	10,803	6086	100.00	89.36	93.71	100.00	<.001
Ashkenazi Jewish	584 (0.49)	373	187	100.00	88.63	93.95	100.00	<.001
Hispanic or Latina	10,196 (8.67)	5382	4393	100.00	91.25	92.75	100.00	<.001
Age group (years)								
18-29	65,352 (55.57)	41,126	21,936	100.00	90.55	94.73	100.00	<.001
30-39	28,541 (24.27)	17,475	9694	100.00	87.60	92.72	100.00	<.001
40-49	1590 (1.35)	9170	5736	100.00	89.35	93.06	100.00	<.001
50-64	7167 (6.09)	3832	3068	100.00	91.99	93.49	100.00	<.001
≥65	945 (0.8)	469	446	100.00	93.70	93.99	100.00	<.001

[a]Sample size excluded women with a history of genetic testing in the high risk category (n=6342), no personal history of cancer (n=7056), or unknown family history (n=19,138). These categories were not mutually exclusive, and 26,062 were excluded from the analysis.

[b]PPV: positive predictive value.

[c]NPV: negative predictive value.

[d]*P* value was calculated using McNemar test.

Table 3. Assess Your Risk (AYR) compared with National Comprehensive Cancer Network (NCCN) criteria plus polycystic ovary syndrome, childhood radiation, and abnormal breast biopsy

	AYR full criteria compared with NCCN criteria							
	Total[a], n (%)	True positive, n	True negative, n	Sensitivity (%)	Specificity (%)	PPV[b] (%)	NPV[c] (%)	*P* value[d]
Overall	117,595 (100)	76,714	40,880	100.00	100.00	100.00	100.00	.32
Race and ethnicity								
White	82,128 (69.83)	55,134	26,994	100.00	100.00	100.00	100.00	—[e]
African American	3886 (3.3)	2277	1608	100.00	99.94	99.96	100.00	.32
Asian	3187 (2.71)	1575	1612	100.00	100.00	100.00	100.00	—
Other	17,614 (14.98)	11,528	6086	100.00	100.00	100.00	100.00	—
Ashkenazi Jewish	584 (0.49)	397	187	100.00	100.00	100.00	100.00	—
Hispanic or Latina	10,196 (8.67)	5803	4393	100.00	100.00	100.00	100.00	—
Age group (years)								
18-29	65,352 (55.57)	43,416	21,936	100.00	100.00	100.00	100.00	—
30-39	28,541 (24.27)	18,846	9694	100.00	99.99	99.99	100.00	.32
40-49	15,590 (13.26)	9854	5736	100.00	100.00	100.00	100.00	—
50-64	7167 (6.09)	4099	3068	100.00	100.00	100.00	100.00	—
≥65[d]	945 (8.03)	499	446	100.00	100.00	100.00	100.00	—

[a]Sample size excluded women with a history of genetic testing in the high risk category (n=6342), no personal history of cancer (n=7056), or unknown family history (n=19,138). These categories were not mutually exclusive, and 26,062 were excluded from the analysis.

[b]PPV: positive predictive value.

[c]NPV: negative predictive value.

[d]*P* value was calculated using McNemar test.

[e]No discordant pairs between groups where all measures are 100.00.

Discussion

Principal Findings

We present the AYR tool as a valuable, new risk assessment tool with great potential to impact breast and ovarian cancer prevention in the general population. Overall, the AYR tool is highly accurate in identifying women at increased risk for breast and ovarian cancer in comparison with the gold standard NCCN criteria. We also identified differences by race, ethnicity, and age group that highlight opportunities for improving the reach of the AYR tool to additional populations as well as objective for future studies to validate the AYR tool among diverse populations. Overall, the Bright Pink AYR tool provides reliable, evidence-based risk assessment to the general as well as clinical populations for the identification of women at increased risk for breast and ovarian cancer.

Existing Risk Assessment Tools Compared With Bright Pink AYR

Bright Pink AYR is unique compared with other web-based risk assessment tools for breast and ovarian cancer, primarily because of its nature as a participant or patient-facing web-based tool designed for use by the general population. This includes, but is not limited to, clinical risk assessment tools such as the Breast Cancer Risk Assessment Tool (BCRAT; also known as the Gail model), the International Breast Cancer Intervention Study (IBIS), the Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA) model, and BCRAPRO [25-29]. There is much overlap between these validated, clinically oriented tools with AYR but also important differences. For example, the BCRAT model uses 7 questions including age, race, ethnicity, health history (age at menarche and history of abnormal breast biopsy), and family history of breast cancer to estimate the 5-year risk and lifetime risk of invasive breast cancer based on the probability of cancer incidence during a defined age range [21,22]. The BCRAT model, however, does not include the risk of genetic mutations but instead refers participants to alternative tools for testing [21,22]. In contrast, although software such as IBIS using the Tyrer-Cuzick model also estimates 5-year and lifetime risk, the model incorporates genetic mutations and breast density into the most recent versions [23,24]. Moreover, both IBIS and BOADICEA were recently updated to include polygenetic risk scores, which is a limitation that other models such as the AYR should examine in future iterations. BRCAPRO is a similar model that relies primarily on basic demographics, family history, and genetic testing results, if available [30]. However, one major difference between AYR and these other tools is that, other than BOADICEA, none include lifestyle or health behavior

risk factors in the models [27]. Overall, although there are many similarities between AYR and existing breast and ovarian cancer clinical risk assessment tools, the major difference is that AYR was designed as a resource for the general population.

Opportunities for Improving Risk Assessment Through Bright Pink AYR

Bright Pink built the AYR for the general population as a tool to help inform women of their risk and to provide education on cancer prevention activities appropriate to their risk level. We believe that the patient-oriented interface, developed with the help of marketing experts, makes AYR a particularly user-friendly risk assessment tool with the potential to reach populations that existing clinical tools may not reach because of issues such as health care access. The Bright Pink AYR tool represents an advance in the field of population-level cancer prevention as a screening tool with the potential to improve the reach and use of web-based breast and ovarian risk assessment among diverse populations.

The Bright Pink AYR tool has also demonstrated a strong reach to potentially underserved populations. Epidemiological data from the Surveillance, Epidemiology, and End Results Program demonstrate disparities in breast and ovarian cancer incidence and mortality by race [31-33]. However, studies show that much of the differences in breast and ovarian cancer incidence and mortality rates by race and ethnicity are driven by factors such as access to recommended screening and treatment [34,35]. To improve early cancer detection among high-risk groups and younger women not yet eligible for cancer screening, it is critical to increase access to and use of risk assessment tools in the general population. The Bright Pink AYR was designed to bridge these gaps. However, studies have demonstrated that other current methods for breast and ovarian cancer risk assessment may not currently be acceptable to diverse populations. Studies by Cragun [36] and Sheppard et al [37] demonstrated the disparities by race in the acceptability and use of breast and ovarian cancer risk assessment. The results demonstrate that breast and ovarian cancer risk assessment in a clinical setting is less likely to be acceptable or used by African American and Hispanic or Latina women compared with White women [36,37]. The observed disparities by race and ethnicity in breast and ovarian cancer risk as well as use of risk assessment tools justify the need to modernize risk assessment to make it more broadly accessible to the general population. We have demonstrated that the Bright Pink AYR reaches these diverse populations in the general population and provides accurate risk assessment; however, the educational follow-up emails also make AYR unique.

The AYR tool is also unique compared with other tools in that health behaviors are collected and used to tailor email messages to promote cancer prevention behaviors. Previous studies have demonstrated the impact of education on health behaviors among high-risk populations. For example, Quach et al [38] examined baseline and 6-month reporting of health behaviors among Ashkenazi Jewish individuals (n=120) who underwent genetic testing for *BRCA1/2*. The study found that women and those with higher education were more likely to report a healthy diet. The results showed no change in diet, vitamin use, or physical activity over time. However, there was no education component for health behaviors incorporated into the genetic counseling sessions. Another study examined health behaviors among participants of self-reported genetic testing (n=3016; with n=136 reporting genetic testing) using data from the Health Information National Trends Survey 4 [39]. The results demonstrated that lifestyle factors were not statistically significantly different between those who reported genetic testing and those who did not. However, the sample size was relatively small, and the study did not examine multivariate models. The AYR tool collects data on health behaviors, unlike many other clinician-oriented risk assessment tools. However, the AYR educational component has a strong potential to impact population-level participation in cancer prevention activities.

Studies have examined the impact of risk assessment programs for breast or ovarian cancer on adherence to cancer prevention guidelines or recommendations for lifestyle risk factors, and most have identified opportunities for improvement [38-43]. For women with a strong family history, Price et al [44] reported that among 748 women in a breast cancer family registry, between baseline and 3-year surveys, 16% were underscreened for mammography and 55% were underscreened for clinical breast examinations. Moreover, a study by Loescher et al [45] examined cancer surveillance behaviors among 107 women aged ≥18 years who presented for genetic risk assessment for breast and ovarian cancer. The results showed that 60% engaged in the minimal level of recommended breast cancer prevention activities, but 70% reported behaviors below optimal guidelines for cancer prevention. The study also found that a lack of physician recommendation was the most commonly reported reason for not engaging in breast or cancer prevention activities [45]. Similarly, Botkin et al [46] studied the impact of *BRCA1* genetic testing on preventative cancer screening behavior over 2 years (n=408). Among women aged ≥40 years, 82% of mutation carriers followed guidelines for screening mammography in the first year and 67% in the second year, which was significantly increased from baseline and greater than the levels observed among noncarriers [46]. These studies provide evidence that knowledge of risk for breast and ovarian cancer may not only impact participation in cancer prevention activities but also that reported levels are not optimal and expanded adoption of combined risk assessment and educational modalities such as the AYR tool may improve cancer prevention in the general population.

There are strengths and limitations of the current analysis. The strengths include the very large sample size of the completions of the AYR tool in a short period and a diverse population. However, although the population was diverse, the frequency by race and ethnicity was lower than expected for some populations. This reflects the opportunity to improve access and use of the AYR tool in diverse populations. Moreover, the questions asked by the Bright Pink AYR tool are categorical and often binomial, such as health or family history and physical activity questions. This limits the level of detailed data collected on high-risk cancers and the ability to evaluate factors such as age at cancer diagnosis and pathology or subtype of breast cancer. Finally, these data were cross-sectional. Longitudinal studies will be necessary to validate the accuracy of AYR risk

assessment in breast and ovarian cancer risk prevention. However, overall, the results demonstrate that AYR accurately classifies women according to breast and ovarian cancer risk using existing gold standard criteria.

These results demonstrate that Bright Pink's AYR accurately classifies women at an increased risk of breast or ovarian cancer. The variation observed by race, ethnicity, and age group demonstrates the need to improve access and use of risk assessment tools in diverse populations. Overall, Bright Pink AYR is a valuable tool for use by the general population as well as patients and clinical providers for early detection and education, which will improve the prevention of breast and ovarian cancer.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplemental tables describing the Bright Pink Assess Your Risk population and characteristics.
[DOCX File , 22 KB - jmir_v24i2e29124_app1.docx]

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin 2016 Jan 07;66(1):7-30 [FREE Full text] [doi: 10.3322/caac.21332] [Medline: 26742998]
2. Daly MB, Pilarski R, Berry M, Buys SS, Farmer M, Friedman S, et al. NCCN guidelines insights: genetic/familial high-risk assessment: breast and ovarian, version 2.2017. J Natl Compr Canc Netw 2017 Jan;15(1):9-20. [doi: 10.6004/jnccn.2017.0003] [Medline: 28040716]
3. DeSantis CE, Bray F, Ferlay J, Lortet-Tieulent J, Anderson BO, Jemal A. International variation in female breast cancer incidence and mortality rates. Cancer Epidemiol Biomarkers Prev 2015 Oct;24(10):1495-1506 [FREE Full text] [doi: 10.1158/1055-9965.EPI-15-0535] [Medline: 26359465]
4. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015 Mar 1;136(5):359-386. [doi: 10.1002/ijc.29210] [Medline: 25220842]
5. Tao Z, Shi A, Lu C, Song T, Zhang Z, Zhao J. Breast cancer: epidemiology and etiology. Cell Biochem Biophys 2015 Jun;72(2):333-338. [doi: 10.1007/s12013-014-0459-6] [Medline: 25543329]
6. Fitzmaurice C. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 29 cancer groups, 2006 to 2016: a systematic analysis for the Global Burden of Disease study. J Clin Oncol 2018 May 20;36(15_suppl):1568. [doi: 10.1200/jco.2018.36.15_suppl.1568]
7. US Preventive Services Task Force, Grossman DC, Curry SJ, Owens DK, Barry MJ, Davidson KW, et al. Screening for ovarian cancer: US preventive services task force recommendation statement. J Am Med Assoc 2018 Feb 13;319(6):588-594. [doi: 10.1001/jama.2017.21926] [Medline: 29450531]
8. Guo F, Hirth JM, Lin Y, Richardson G, Levine L, Berenson AB, et al. Use of BRCA Mutation Test in the U.S., 2004-2014. Am J Prev Med 2017 Jun;52(6):702-709 [FREE Full text] [doi: 10.1016/j.amepre.2017.01.027] [Medline: 28342662]
9. Oosterwijk JC, de Vries J, Mourits MJ, de Bock GH. Genetic testing and familial implications in breast-ovarian cancer families. Maturitas 2014 Aug;78(4):252-257. [doi: 10.1016/j.maturitas.2014.05.002] [Medline: 24894332]
10. Afghahi A, Kurian AW. The changing landscape of genetic testing for inherited breast cancer predisposition. Curr Treat Options Oncol 2017 May;18(5):27. [doi: 10.1007/s11864-017-0468-y] [Medline: 28439798]
11. Kuchenbaecker KB, Hopper JL, Barnes DR, Phillips K, Mooij TM, Roos-Blom M, BRCA1 and BRCA2 Cohort Consortium, et al. Risks of breast, ovarian, and contralateral breast cancer for BRCA1 and BRCA2 mutation carriers. J Am Med Assoc 2017 Jun 20;317(23):2402-2416. [doi: 10.1001/jama.2017.7112] [Medline: 28632866]
12. Chen S, Iversen ES, Friebel T, Finkelstein D, Weber BL, Eisen A, et al. Characterization of BRCA1 and BRCA2 mutations in a large United States sample. J Clin Oncol 2006 Feb 20;24(6):863-871 [FREE Full text] [doi: 10.1200/JCO.2005.03.6772] [Medline: 16484695]
13. Hoskins KF, Zwaagstra A, Ranz M. Validation of a tool for identifying women at high risk for hereditary breast cancer in population-based screening. Cancer 2006 Oct 15;107(8):1769-1776 [FREE Full text] [doi: 10.1002/cncr.22202] [Medline: 16967460]
14. Rockhill B, Spiegelman D, Byrne C, Hunter DJ, Colditz GA. Validation of the Gail et al. model of breast cancer risk prediction and implications for chemoprevention. J Natl Cancer Inst 2001 Mar 07;93(5):358-366. [doi: 10.1093/jnci/93.5.358] [Medline: 11238697]
15. Shiovitz S, Korde LA. Genetics of breast cancer: a topic in evolution. Ann Oncol 2015 Jul;26(7):1291-1299 [FREE Full text] [doi: 10.1093/annonc/mdv022] [Medline: 25605744]

16. Crawford B, Adams SB, Sittler T, van den Akker J, Chan S, Leitner O, et al. Multi-gene panel testing for hereditary cancer predisposition in unsolved high-risk breast and ovarian cancer patients. Breast Cancer Res Treat 2017 Jun 9;163(2):383-390 [FREE Full text] [doi: 10.1007/s10549-017-4181-0] [Medline: 28281021]
17. Daly MB, Pilarski R, Axilbund JE, Berry M, Buys SS, Crawford B, et al. Genetic/familial high-risk assessment: breast and ovarian, version 2.2015. J Natl Compr Canc Netw 2016 Feb;14(2):153-162. [doi: 10.6004/jnccn.2016.0018] [Medline: 26850485]
18. Bjørnslett M, Dahl AA, Sørebø O, Dørum A. Psychological distress related to BRCA testing in ovarian cancer patients. Fam Cancer 2015 Dec;14(4):495-504. [doi: 10.1007/s10689-015-9811-2] [Medline: 25980896]
19. Graves KD, Vegella P, Poggi EA, Peshkin BN, Tong A, Isaacs C, et al. Long-term psychosocial outcomes of BRCA1/BRCA2 testing: differences across affected status and risk-reducing surgery choice. Cancer Epidemiol Biomarkers Prev 2012 Mar;21(3):445-455 [FREE Full text] [doi: 10.1158/1055-9965.EPI-11-0991] [Medline: 22328347]
20. Cortez S, Milbrandt M, Kaphingst K, James A, Colditz G. The readability of online breast cancer risk assessment tools. Breast Cancer Res Treat 2015 Nov;154(1):191-199 [FREE Full text] [doi: 10.1007/s10549-015-3601-2] [Medline: 26475705]
21. Costantino JP, Gail MH, Pee D, Anderson S, Redmond CK, Benichou J, et al. Validation studies for models projecting the risk of invasive and total breast cancer incidence. J Natl Cancer Inst 1999 Sep 15;91(18):1541-1548. [doi: 10.1093/jnci/91.18.1541] [Medline: 10491430]
22. Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst 1989 Dec 20;81(24):1879-1886. [doi: 10.1093/jnci/81.24.1879] [Medline: 2593165]
23. Brentnall AR, Harkness EF, Astley SM, Donnelly LS, Stavrinos P, Sampson S, et al. Mammographic density adds accuracy to both the Tyrer-Cuzick and Gail breast cancer risk models in a prospective UK screening cohort. Breast Cancer Res 2015 Dec 01;17(1):147 [FREE Full text] [doi: 10.1186/s13058-015-0653-5] [Medline: 26627479]
24. Brentnall AR, Cuzick J, Buist DS, Bowles EJ. Long-term accuracy of breast cancer risk assessment combining classic risk factors and breast density. JAMA Oncol 2018 Sep 01;4(9):e180174 [FREE Full text] [doi: 10.1001/jamaoncol.2018.0174] [Medline: 29621362]
25. Daly MB, Axilbund JE, Buys S, Crawford B, Farrell CD, Friedman S, National Comprehensive Cancer Network. Genetic/familial high-risk assessment: breast and ovarian. J Natl Compr Canc Netw 2010 May;8(5):562-594. [doi: 10.6004/jnccn.2010.0043] [Medline: 20495085]
26. Nelson H, Huffman L, Fu R, Harris E, Walker M, Bougatsos C. Genetic risk assessment and BRCA mutation testing for breast and ovarian cancer susceptibility. In: U.S. Preventive Services Task Force Evidence Syntheses, formerly Systematic Evidence Reviews. Rockville (MD): Agency for Healthcare Research and Quality (US); 2005.
27. Terry M, Liao Y, Whittemore A, Leoce N, Buchsbaum R, Zeinomar N. 10-year performance of four models of breast cancer risk: a validation study. Lancet Oncol 2019;20(4):504-517 [FREE Full text] [doi: 10.20900/mo.20190017]
28. Choudhury PP, Brook MN, Hurson AN, Lee A, Mulder CV, Coulson P, et al. Comparative validation of the BOADICEA and Tyrer-Cuzick breast cancer risk models incorporating classical risk factors and polygenic risk in a population-based prospective cohort of women of European ancestry. Breast Cancer Res 2021 Feb 15;23(1):22 [FREE Full text] [doi: 10.1186/s13058-021-01399-7] [Medline: 33588869]
29. Choudhury PP, Wilcox AN, Brook MN, Zhang Y, Ahearn T, Orr N, et al. Comparative validation of breast cancer risk prediction models and projections for future risk stratification. J Natl Cancer Inst 2020 Mar 01;112(3):278-285 [FREE Full text] [doi: 10.1093/jnci/djz113] [Medline: 31165158]
30. Berry DA, Iversen ES, Gudbjartsson DF, Hiller EH, Garber JE, Peshkin BN, et al. BRCAPRO validation, sensitivity of genetic testing of BRCA1/BRCA2, and prevalence of other breast cancer susceptibility genes. J Clin Oncol 2002 Jun 01;20(11):2701-2712. [doi: 10.1200/jco.2002.05.121]
31. Hung M, Ekwueme DU, Rim SH, White A. Racial/ethnicity disparities in invasive breast cancer among younger and older women: an analysis using multiple measures of population health. Cancer Epidemiol 2016 Dec;45:112-118 [FREE Full text] [doi: 10.1016/j.canep.2016.10.013] [Medline: 27792934]
32. DeSantis CE, Fedewa SA, Goding SA, Kramer JL, Smith RA, Jemal A. Breast cancer statistics, 2015: convergence of incidence rates between black and white women. CA Cancer J Clin 2015 Oct 29 [FREE Full text] [doi: 10.3322/caac.21320] [Medline: 26513636]
33. Edwards BK, Noone A, Mariotto AB, Simard EP, Boscoe FP, Henley SJ, et al. Annual Report to the Nation on the status of cancer, 1975-2010, featuring prevalence of comorbidity and impact on survival among persons with lung, colorectal, breast, or prostate cancer. Cancer 2014 May 01;120(9):1290-1314 [FREE Full text] [doi: 10.1002/cncr.28509] [Medline: 24343171]
34. Bandera EV, Lee VS, Rodriguez-Rodriguez L, Powell CB, Kushi LH. Racial/ethnic disparities in ovarian cancer treatment and survival. Clin Cancer Res 2016 Dec 01;22(23):5909-5914 [FREE Full text] [doi: 10.1158/1078-0432.CCR-16-1119] [Medline: 27521449]
35. Freeman HP. Patient navigation: a community based strategy to reduce cancer disparities. J Urban Health 2006 Mar;83(2):139-141 [FREE Full text] [doi: 10.1007/s11524-006-9030-0] [Medline: 16736361]

36. Cragun D, Weidner A, Lewis C, Bonner D, Kim J, Vadaparampil ST, et al. Racial disparities in BRCA testing and cancer risk management across a population-based sample of young breast cancer survivors. Cancer 2017 Jul 01;123(13):2497-2505 [FREE Full text] [doi: 10.1002/cncr.30621] [Medline: 28182268]
37. Sheppard VB, Mays D, LaVeist T, Tercyak KP. Medical mistrust influences black women's level of engagement in BRCA 1/2 genetic counseling and testing. J Natl Med Assoc 2013;105(1):17-22 [FREE Full text] [doi: 10.1016/s0027-9684(15)30081-x] [Medline: 23862292]
38. Quach J, Porter K, Leventhal H, Kelly KM. Health behaviors among Ashkenazi Jewish individuals receiving counseling for BRCA1 and BRCA2 mutations. Fam Cancer 2009;8(3):241-250. [doi: 10.1007/s10689-009-9235-y] [Medline: 19184534]
39. Quillin JM. Lifestyle risk factors among people who have had cancer genetic testing. J Genet Couns 2016 Oct;25(5):957-964. [doi: 10.1007/s10897-015-9925-6] [Medline: 26659117]
40. Forman A, Hall M. Influence of race/ethnicity on genetic counseling and testing for hereditary breast and ovarian cancer. Breast J 2009;15 Suppl 1(s1):56-62. [doi: 10.1111/j.1524-4741.2009.00798.x] [Medline: 19775331]
41. Simon MS, Petrucelli N. Hereditary breast and ovarian cancer syndrome : the impact of race on uptake of genetic counseling and testing. Methods Mol Biol 2009;471:487-500. [doi: 10.1007/978-1-59745-416-2_25] [Medline: 19109796]
42. Shulman LP. Hereditary breast and ovarian cancer (HBOC): clinical features and counseling for BRCA1 and BRCA2, Lynch syndrome, Cowden syndrome, and Li-Fraumeni syndrome. Obstet Gynecol Clin North Am 2010 Mar;37(1):109-133. [doi: 10.1016/j.ogc.2010.03.003] [Medline: 20494261]
43. Kurian AW. BRCA1 and BRCA2 mutations across race and ethnicity: distribution and clinical implications. Curr Opin Obstet Gynecol 2010 Mar;22(1):72-78. [doi: 10.1097/GCO.0b013e328332dca3] [Medline: 19841585]
44. Price MA, Butow PN, Charles M, Bullen T, Meiser B, McKinley JM, kConFab PsychosocialClinical Follow-Up Groups, kConFab Investigators. Predictors of breast cancer screening behavior in women with a strong family history of the disease. Breast Cancer Res Treat 2010 Nov;124(2):509-519. [doi: 10.1007/s10549-010-0868-1] [Medline: 20364401]
45. Loescher LJ, Lim KH, Leitner O, Ray J, D'Souza J, Armstrong CM. Cancer surveillance behaviors in women presenting for clinical BRCA genetic susceptibility testing. Oncol Nurs Forum 2009 Mar;36(2):57-67. [doi: 10.1188/09.onf.e57-e67] [Medline: 19273395]
46. Botkin JR, Smith KR, Croyle RT, Baty BJ, Wylie JE, Dutson D, et al. Genetic testing for a BRCA1 mutation: prophylactic surgery and screening behavior in women 2 years post testing. Am J Med Genet A 2003 Apr 30;118A(3):201-209. [doi: 10.1002/ajmg.a.10102] [Medline: 12673648]

Abbreviations

AYR: Assess Your Risk
BCRAT: Breast Cancer Risk Assessment Tool
BOADICEA: Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm
IBIS: International Breast Cancer Intervention Study
NCCN: National Comprehensive Cancer Network
NPV: negative predictive value
PCOS: polycystic ovary syndrome
PPV: positive predictive value

Edited by R Kukafka; submitted 26.03.21; peer-reviewed by K Crew, K Shannon; comments to author 11.05.21; revised version received 23.07.21; accepted 02.12.21; published 24.02.22.

Please cite as:
Hibler EA, Fought AJ, Kershaw KN, Molsberry R, Nowakowski V, Lindner D
Novel Interactive Tool for Breast and Ovarian Cancer Risk Assessment (Bright Pink Assess Your Risk): Development and Usability Study
J Med Internet Res 2022;24(2):e29124
URL: https://www.jmir.org/2022/2/e29124
doi:10.2196/29124
PMID:35200148

Original Paper

Health Information Systems for Older Persons in Select Government Tertiary Hospitals and Health Centers in the Philippines: Cross-sectional Study

Angely P Garcia[1*], MPH, RN; Shelley F De La Vega[1,2*], MD, MSc; Susan P Mercado[3*], MD, MPH

[1]Institute on Aging, National Institutes of Health, University of the Philippines Manila, Manila, Philippines

[2]College of Medicine, University of the Philippines Manila, Manila, Philippines

[3]National Telehealth Center, National Institutes of Health, University of the Philippines Manila, Manila, Philippines

[*]all authors contributed equally

Corresponding Author:
Angely P Garcia, MPH, RN
Institute on Aging
National Institutes of Health
University of the Philippines Manila
Rm 211 NIH Bldg
623 Pedro Gil St. Ermita
Manila, 1000
Philippines
Phone: 63 9064029690
Fax: 63 85264349
Email: apgarcia@up.edu.ph

Abstract

Background: The rapid aging of the world's population requires systems that support health facilities' provision of integrated care at multiple levels of the health care system. The use of health information systems (HISs) at the point of care has shown positive effects on clinical processes and patient health in several settings of care.

Objective: We sought to describe HISs for older persons (OPs) in select government tertiary hospitals and health centers in the Philippines. Specifically, we aimed to review the existing policies and guidelines related to HISs for OPs in the country, determine the proportion of select government hospitals and health centers with existing health information specific for OPs, and describe the challenges related to HISs in select health facilities.

Methods: We utilized the data derived from the findings of the Focused Interventions for Frail Older Adults Research and Development Project (FITforFrail), a cross-sectional and ethics committee–approved study. A facility-based listing of services and human resources specific to geriatric patients was conducted in purposively sampled 27 tertiary government hospitals identified as geriatric centers and 16 health centers across all regions in the Philippines. We also reviewed the existing policies and guidelines related to HISs for OPs in the country.

Results: Based on the existing guidelines, multiple agencies were involved in the provision of services for OPs, with several records containing health information of OPs. However, there is no existing HIS specific for OPs in the country. Only 14 (52%) of the 27 hospitals and 4 (25%) of the 16 health centers conduct comprehensive geriatric assessment (CGA). All tertiary hospitals and health centers are able to maintain medical records of their patients, and almost all (26/27, 96%) hospitals and all (16/16, 100%) health centers have data on top causes of morbidity and mortality. Meanwhile, the presence of specific disease registries varied per hospitals and health centers. Challenges to HISs include the inability to update databases due to inadequately trained personnel, use of an offline facility–based HIS, an unstable internet connection, and technical issues and nonuniform reporting of categories for age group classification.

Conclusions: Current HISs for OPs are characterized by fragmentation, multiple sources, and inaccessibility. Barriers to achieving appropriate HISs for OPs include the inability to update HISs in hospitals and health centers and a lack of standardization by age group and disease classification. Thus, we recommend a 1-person, 1-record electronic medical record system for OPs and the disaggregation and analysis across demographic and socioeconomic parameters to inform policies and programs that address

the complex needs of OPs. CGA as a required routine procedure for all OPs and its integration with the existing HISs in the country are also recommended.

(J Med Internet Res 2022;24(2):e29541) doi:10.2196/29541

KEYWORDS

health information systems; the Philippines; aged; hospitals; community health centers; database; geriatric assessment; elderly; digital health; medical records; health policy

Introduction

The world's population is rapidly aging, from the 12% estimate in 2015 to the 22% total global population in 2050 [1]. In the Philippines, 7.5 million, or 7.5%, of the total country population in 2015 were senior citizens (aged 60 years and above) [2]. Recognizing their complex health needs and considering that sound and reliable information is the foundation of decision making across all health systems, the World Health Organization (WHO) developed the Global Strategy and Action Plan on Aging and Health (GSAP 2016-2020), which includes adapting information systems to collect, analyze, and report data on intrinsic trends in the capacity of the aging population [3].

Comprehensive geriatric assessment (CGA) is a form of collecting, analyzing, and reporting data on the intrinsic capacity of an older person (OP). It is a multidimensional, multidisciplinary diagnostic and treatment process conducted by a team of health professionals through a systematic evaluation that identifies a variety of treatable health problems and leads to better health outcomes [4]. It is currently being utilized in different settings, government and private facilities, outpatient and inpatient care, primary care, and research. It contains multiple data points and essential health information about OPs that must be considered in providing holistic and integrated care. Based on findings of meta-analyses [5-10], CGA leads to improved detection and documentation of geriatric problems as well as improvement of health outcomes, such as improvement of functional status, prevention of hospitalization, and reduction in readmission rates or mortality, depending on the specific model and setting in which it is implemented [4]. Furthermore, recent evidence on the cost and effects of CGA showed a reduction in the need for hospital care days in a high-risk population of older adults, which could be of great importance in managing the increasing prevalence of frailty and multimorbidity [11]. This information is also crucial for program and policy development.

One of the main challenges of today's health system in the country is access to real-time information for decision making [12]. The 2018 Philippine's health system review highlighted that integrating and harmonizing all existing health-related information systems and data sources, and the inadequacy of a governance structure on information and communication technologies (ICT) are critical challenges [13]. Moreover, the privacy of heath information was also identified as a challenge in policy and practice [14].

The rapid aging of the population requires systems that support health facilities' provision of integrated care at multiple levels of the health care system. A health information system (HIS) that maintains "1 person, 1 record" facilitates efficient provision of services for OPs. Furthermore, the use of HISs at the point of care has shown positive impacts on clinical processes and patient health in multiple settings of care [15]. The adoption of health information exchange (HIE) programs has proven to lessen utilization of health care services, such as ambulatory care and hospital readmissions, and allow smooth transition from inpatient to outpatient care [16,17].

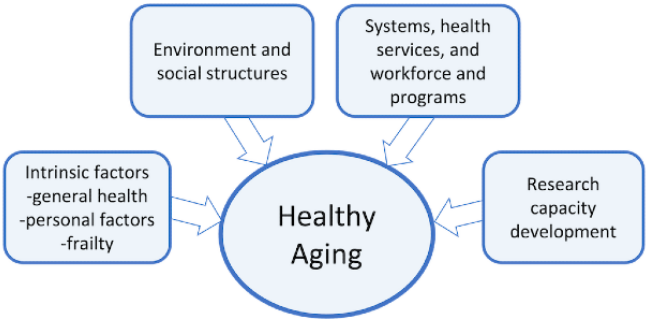
In recognition of the need for Filipino senior citizens to receive appropriate geriatric health care services, the Department of Health (DOH) provided funding for upgrading the 27 DOH-retained hospitals across regions where geriatric centers will be established [18]. CGA will be conducted in these centers and in primary care settings through Guidelines on the Adoption of Baseline Primary Health Care Guarantees for All Filipinos (DOH Administrative Order [AO] no. 2017-002) [19].

Given the rapid aging population, complex needs of OPs, importance of health information in the delivery of services, and challenges to health information in general, identifying the current status of HISs for OPs is significant in aligning the health system in the country to achieve healthy aging. This is especially true for government tertiary hospitals and health centers where OPs usually access health care.

Figure 1 shows the Focused Interventions for Frail Older Adults Research and Development Project (FITforFrail) framework adapted from the WHO Healthy Aging Framework, which defines healthy aging as the process of developing and maintaining the functional ability that enables well-being in older age [20]. The systems, health services, workforce and programs, intrinsic factors, environment and social structures, and research capacity development are essential parts of the whole-of-system approach that supports healthy aging.

Since healthy aging is the main focus of the GSAP, wherein 1 of the key strategies is aligning health systems to the needs of OPs [3], FITforFrail Study 1 concentrated on the analysis of current health systems for aging in the Philippines. The systems, health services, and workforce and programs, as well as aspects of the environment and social structures, were covered by Study 1, where mixed methods of data collection were utilized.

Figure 1. FITforFrail healthy aging framework. FITforFrail: Focused Interventions for Frail Older Adults Research and Development Project.



According to WHO, health systems need to be transformed and realigned to ensure access to evidence-based health interventions responsive to the needs of OPs [3]. A HIS is one of the building blocks of the health system. It provides the underpinnings for decision making where data generation, compilation, analysis and synthesis, and communication and use are its key functions [21]. With the advent of technology, eHealth or the use of ICT for health can maximize its potential toward integrated care of OPs.

The HIS is particularly under the systems, health services, and workforce and programs wherein review of policies related to OPs and listing of services and workforce specific for geriatric patients are conducted.

The data on HISs in general and even among specific population groups in the Philippines are limited. Thus, this paper sought to describe the existing HISs specific for OPs, especially among government tertiary hospitals and health centers across regions in the country. Specifically, it aimed to review the existing policies and guidelines related to HISs for OPs in the Philippines; determine the proportion of select government hospitals and health centers with existing health information specific for OPs, such as CGA, medical records, analysis of top causes of morbidity and mortality, and registry of specific diseases, and electronic medical records (EMRs); and describe the challenges related to HISs in select tertiary hospitals and health centers.

Methods

FITForFrail

FITforFrail is a research and development project funded by the DOH through the Philippine Council for Health Research and Development (PCHRD) under the Advancing Health through Evidence-Assisted Decisions with Health Policy Systems Research (AHEAD-HPSR) program. Using the WHO Healthy Aging Framework, the project aims to identify the current health system for the aging population and describe the health status of OPs in select communities. FITforFrail Study 1 analyzed the health system, and FITforFrail Study 2 evaluated the health status of OPs, with a focus on frailty.

Study Design

A cross-sectional research study design using mixed methods of data collection was utilized. Mixed methods and community participation were hallmarks of this research. For this specific paper, a review of policies and papers related to OPs and a facility-based listing of health services and workforce specific for OPs were conducted to collect the data on HISs.

Sampling

Purposive sampling was used in selecting study sites. The inclusion criteria for the hospitals were (1) Ministry of Health–/DOH-retained government tertiary hospitals and (2) geriatric centers identified through the Philippine Health Development Plan 2017-2022. For health centers, they had to be within the catchment area of the identified hospital in the region. All the 27 hospitals identified as geriatric centers and 17 health centers within the catchment area across all regions were included in the study.

Study Setting and Participants

The researchers conducted a listing of health services and workforce specific for OPs in the 27 hospitals identified as geriatric centers and the 17 health centers within their catchment area. An advance copy of the listing tool along with the letter addressed to the heads of the institutions was sent prior to actual data collection. The heads of the institutions assigned and identified focal or point persons to be interviewed to provide their facility data. These identified point persons served as key informants. They were mostly in charge of the geriatric program in their institutions. The research team scheduled separate meetings with the informants to explain the study and obtain consent prior to actual data collection. Policies and the existing literature on HISs were also reviewed.

Ethics Clearance

FITforFrail obtained a total of 6 ethics approvals from the University of the Philippines Manila Research Ethics Board (UPMREB), the Single Joint Research Ethics Board (SJREB), and 4 institutional review boards of hospitals (Multimedia Appendix 1). The UPMREB oversight applies to UP Manila researchers and non-UP Manila researchers doing research in non-UP Manila sites with no local ethics review committee (as mandated by the Philippine Health Research Ethics Board [PHREB]), while the SJREB is a joint review mechanism among the PHREB–duly accredited research ethics committees (RECs) of DOH hospitals. The rest of the reviews and approvals were from the DOH hospitals that required separate institutional review.

Data Collection

Desk Review

Policies were collected through consultation, online search or bibliographic databases visits, and manual search or onsite library visits. For bibliographic databases and online search, the following search terms were used covering the period of 1980 to July 2020: “aging,” “senior citizens,” “older persons,” “Philippines,” “Republic Act,” “memorandum,” “circulars,” “policy,” “administrative order,” “health information system,” “information systems,” and “programs.”

Listing of Services and Workforce

A facility-based listing of services and workforce specific for OPs was conducted. The DOH hospitals identified as geriatric centers were selected as study sites. For primary care units, health centers within the catchment area of the identified regional hospitals were purposely selected across 17 regions in the country. A total of 27 DOH-retained hospitals identified as geriatric centers and 17 health centers were visited for the facility-based listing of services and workforce, with particular attention to HISs for OPs.

A checklist or facility-based listing form was used (Multimedia Appendix 2). The listing form was developed by trained research assistants through policy review and series of consultation meetings with the project and study leaders. The sections of the listing form are as follows: facility demographics, human resources, competencies and training, health services, health financing, information system, and health policies and programs. The specific section on HISs contains questions on patient medical records, disease registries, online databases, and reasons for not having such registries and databases. Moreover, the question on CGA was included under the Health Services section.

Data Management and Analysis

The collected data from the listing were entered in a password-protected EpiInfo data entry program. Data verification and cleaning were conducted using Microsoft Excel through the help of a statistical assistant and under the supervision of a statistician. Cleaned data sets were endorsed to the statistician for analysis. Descriptive statistics (means, SDs, and frequency distribution) were calculated for all continuous and categorical variables measured using Stata (StataCorp).

Results

Policies and Guidelines on Health Information Systems for Older Persons

Republic Act No. 11223 or the Universal Health Care (UHC) Act of 2019 state that all health facilities are required to maintain a HIS consistent with DOH standards, which will be electronically uploaded on a regular basis through interoperable systems [22]. The DOH and PhilHealth will fund and manage the development, quality assurance, and maintenance of the information systems. Under the implementation of the UHC Act is the establishment of a HIS in every health facility, which requires multiple key players for the provision of population- and individual-based health services, including the services for OPs. The DOH, PhilHealth, and the Department of Interior and Local Government (DILG) will integrate all local health systems into a province-wide health system. The private sector will also be encouraged to participate in the integrated local health system through a contractual arrangement.

Prior to the UHC Act, the DOH issued standard policies, procedures, and guidelines governing all ICT-related work in 2005 [13]. It also established the Knowledge Management and Information Technology Service (KMITS); developed the Department of Health Enterprise Architecture (DOH EA) for HISs, which is national in scope; implemented information systems using client-server technology; and established an eHealth framework [23]. A part of the eHealth framework is the Philippine Health Information Exchange (PHIE) through the Joint DOH-DOST-PhilHealth AO no. 2016-001. It aims to achieve integrated health care services and delivery that are also seamlessly responsive, efficient, cost effective, and in real time [24].

The Joint AO no. 2016-003 of DOH and PhilHealth gave way to the adoption of PHIE Lite, which aims to institutionalize the implementation of a harmonized approach and system in developing applications and information systems [25]. OPs were included in the initial priorities of PHIE Lite interoperability as they are included as expanded primary benefit care (ePCB)-entitled sponsored members.

The National Health Data Dictionary (NHDD) and the Unified Health Management Information System (UHMIS), and interoperability standards were also developed and implemented through DOH AO nos. 2013-025 and 2015-037 [26,27]. Unfortunately, the latest version of the NHDD (version 2.0) [28] do not include standard age group classification (young, middle, and oldest old) and relevant diseases, such as geriatric syndromes (ie, dementia, frailty, malnutrition, polypharmacy, and incontinence).

A program dedicated to OPs, the National Health and Wellness Program for Senior Citizens (NHWPSC) through the DOH AO no. 2015-009, was established. One of its objectives is to establish and maintain a database management system and conduct research in the development of evidence-based policies for senior citizens [29]. To date, there is no database management system specific for OPs.

To summarize, multiple agencies are involved in the provision of services for OPs, with several records containing health information about OPs. Moreover, there is no system to integrate or enable interoperability of data systems of OPs at primary, secondary, or tertiary levels of care. Hence, a provider for an OP would be unable to access medical, social, or insurance information in a single record.

Health Information Specific for Older Persons

Table 1 summarizes the health information for OPs in visited government tertiary hospitals and health centers across all regions in the country.

Table 1. Health information for OPs[a] in government tertiary hospitals and health centers, 2019-2020.

Health information	Hospitals (N=27), n (%)	Health centers (N=16), n (%)
Facilities	27 (100)	16 (100)
CGA[b]	14 (52)	4 (25)
Medical records of patients	27 (100)	16 (100)
Data on top causes of mortality and morbidity	26 (96)	16 (100)
Registry of diseases of OPs	20 (74)	13 (81)
Diseases in the registry		
Hypertension	18 (67)	13 (81)
Diabetes mellitus	18 (67)	13 (81)
CVD[c]	18 (67)	10 (62)
Stroke or cerebrovascular attack	18 (67)	10 (62)
Heart attack/myocardial infarction	17 (63)	10 (62)
Respiratory tract diseases	20 (74)	13 (81)
Cancer	20 (74)	5 (31)
Mental disorders	7 (26)	3 (19)
Disability	9 (33)	7 (44)
Online web-based database		
Patient records	23 (85)	10 (62)
iHoMIS[d]	15 (56)	0
UDRS[e]	10 (37)	1 (6)
iClinicSys[f]	0	10 (62)
Others (Bizbox, MedSys, Medix, CHITS[g])	8 (30)[h]	1 (6)[i]
Not updated regularly	4 (15)	6 (37)
Reasons[j]		
No trained/not enough personnel	4 (15)	2 (12)
Unstable internet	0	2 (12)
Use of an offline system	2 (7)	1 (6)
Technical issues	0	2 (12)

[a]OP: older person.

[b]CGA: comprehensive geriatric assessment.

[c]CVD: cardiovascular disease.

[d]iHOMIS: Integrated Hospital Operations and Management Information System.

[e]UDRS: Unified Disease Registry System.

[f]iClinicSys: Integrated Clinic Information System.

[g]CHITS: Community Health Information Tracking System.

[h]Values for Bizbox, MedSys, and Medix.

[i]Value for CHITS.

[j]Multiple responses possible.

Comprehensive Geriatric Assessment

A total of 27 DOH tertiary hospitals and 17 health centers were visited. Of the 17 health centers, only 16 (94%) have facility-based listing data. There was no information obtained from a health center in Region IV-B, a cluster of islands in southern Luzon, Philippines [30]. The specific question on CGA was in the Health Service Delivery section of the checklist.

The study revealed that only 14 (52%) of the 27 hospitals identified as geriatric centers conduct CGA for their geriatric patients (Table 1). Of these, only 5 (18%) hospitals use CGA to screen for all their geriatric patients; the rest have specific

conditions or guidelines regarding to whom they can administer CGA. Most hospitals would only utilize CGA in specific age brackets; other hospitals would only do so through referrals, when the patient is admitted, or when they think the patient is frail or at risk. Commonly reported reasons for not administering CGA to all OPs in hospitals include the lack of manpower, inadequate trained personnel, and the length of the assessment. However, of the 16 health centers, only 4 (25%) conduct CGA for their geriatric patients.

Medical Records and Registries for OPs in Hospitals and Health Centers

All 27 hospitals and 16 health centers maintain medical records of their patients. The data on the top causes of mortality are available in almost all (26/27, 96%) visited hospitals and all (16/16, 100%) health centers. When asked whether the facilities have a registry of diseases of OPs, there are more health centers than hospitals that have these (13/16 [81%] vs 20/27 [74%]), as summarized in Table 1.

In terms of specific registries (Table 1), hospitals have better registries on cardiovascular disease (CVD; 18/27 [67%] vs 10/16 [62%]), stroke (18/27 [67%] vs 10/16 [62%]), heart attack (17/27 [63%] vs 10/16 [62%]), cancer (20/27 [74%] vs 5/16 [31%]), and mental disorders (7/27 [26%] vs 3/16 [19%]). However, health centers have better registries on hypertension (13/16 [81%] vs 18/27 [67%]), diabetes (13/16 [81%] vs 18/27 [67%]), respiratory tract diseases (13/16 [81%] vs 20/27 [74%]), and disability (7/16 [44%] vs 9/27 [33%]). Whether these registries are or are not CGA based is not known, as this was not covered by the study and was considered 1 of its limitations.

There are more hospitals that utilize online web-based database of patients records than health centers (23/27 [85%] vs 10/16 [62%]). More than half (15/27, 56%) of the hospitals utilize the Integrated Hospital Operations and Management Information (iHOMIS), and more than a quarter (10/27, 37%) utilize the Unified Disease Registry System (UDRS). iHOMIS is a Windows-based computerized hospital information system for government hospitals, while the UDRS is a unified registry that contains an injury surveillance system, an integrated noncommunicable diseases registry, a violence against children and women registry, and a persons with disabilities registry [31].

Other third-party providers, such as BizBox, MedSys, and Medix, were also reported. Bizbox is a PhilHealth-accredited health information technology provider that passed the eClaims certification on the case rate system [32]. The MedSys EMR is a web-based application developed for physicians and staff within a health care institution to ensure accuracy, privacy, and service efficiency [33]. Lastly, Medix is a cloud-based clinic management software that helps practitioners improve their clinic operations [34]. Some of these are also being utilized by government hospitals despite the availability of DOH-maintained iHOMIS.

For the health centers, more than half (10/16, 62%) utilize an online web-based database for patient records through the Integrated Clinic Information System (iClinicSys), while only 1 (6%) uses the Community Health Information Tracking System (CHITS), as shown in Table 1. iClinicSys is a system owned by the DOH that efficiently and effectively monitors patient cases in rural health units (RHUs) [31], while CHITS is an EMR system for government primary care health centers in the Philippines [35].

Challenges Related to HISs in Select Tertiary Hospitals and Health Centers

In terms of management of HISs, the most common reasons for not regularly updating the web-based database are a lack of or inadequate trained personnel to maintain and manage the information systems (in 4/27 [15%] of hospitals and 2/16 [12%] of health centers), an unstable internet connection (2/16 [12%] of health centers), the use of an offline system (1/16 [6%] of health centers), and technical issues (2/16 [12%] of health centers), as shown in Table 1.

Discussion

Principal Findings

This study described HISs specific for OPs, especially among government tertiary hospitals and health centers across regions in the Philippines. It reviewed the existing policies and guidelines and determined the proportion of select government hospitals and health centers with existing health information specific for OPs, such as CGA, medical records, top causes of morbidity and mortality, registries of specific diseases, and EMRs. Furthermore, challenges related to HISs in select health facilities were described.

There are various HISs in the country. For primary care benefit providers, the following are the DOH-accredited EMR systems: iClinicSys, CHITS, Segworks Tecknologies (Seg-RHIS), the eHatid local government unit (LGU), Secure Health Information Network and Exchange (SHINE OS+), and Wireless Access for Health (WAH) [36]. Furthermore, the DOH maintains 10 information systems and databases. These include the Electronic Drug Price Monitoring System (EDPMS), iClinicSys, the Integrated Chronic Non-Communicable Disease Registry System (ICNCDRS), the Integrated Drug Test Operations Management Information System (IDTOMIS), iHOMIS, the Integrated TB Information System (ITIS), the Online National Electronic Injury Surveillance System (ONEISS), the Philippine Registry for Persons with Disabilities (PRPWD), the National Rabies Information System (NaRIS), and the Violence Against Women and Children Registry System (VAWCRS) [31]. In addition, there are other private or third-party providers of HISs in the country, such as BizBox, MedSys, and Medix.

Among the existing HISs maintained by the DOH, there is no specific one for OPs. The data on OPs can be distributed in almost all existing HISs (ie, PRPWD; ICNCDRS; online reporting of cancer, diabetes, chronic obstructive pulmonary disease, stroke, blindness, mental, coronary artery disease, and renal data from health facilities; ITIS; ONEISS; and other HISs). All these systems require log-in credentials; thus, only authorized personnel have access.

Based on the policies and literature review, there are policies and guidelines that support the establishment and integration

of HISs for OPs. However, there is no current database management system specific for OPs to date, and the data from the existing HISs maintained by the DOH are not readily accessible. Geriatric syndromes, including frailty, malnutrition, dementia, incontinence, and polypharmacy, are not in the NHDD.

There are multiple information systems and agencies involved in the provision of services and sources of health information about OPs, which leads to fragmented health information about OPs in the country. Given the limited accessibility and fragmentation, coming up with evidence for program and policy development that will address the needs of OPs is a major challenge.

More than half of the hospitals identified as geriatric centers and only a quarter of the health centers conduct CGA for their geriatric patients. According to the DOH AO no. 2017-001, "All older patients with a positive risk screen should have a Comprehensive Interdisciplinary Geriatric Assessment for individual special complex needs" and the "Comprehensive Geriatric Assessment should be updated prior to discharge in chronic care facilities and made available to accepting facilities or carers and vice versa" [37].

This study found a limitation in the conduct of CGA, especially in the primary care setting. Not all visited DOH hospitals, although being identified as geriatric centers, conduct CGA. The commonly reported reasons for not administering CGA to all OPs in hospitals include the lack of personnel, inadequate trained personnel, and the length of the assessment.

All visited hospitals and health centers maintain medical records of their patients. The data on the top causes of mortality are available in all health centers and almost all visited hospitals. There are more hospitals that utilize online web-based databases of patients records than health centers. More than half of the hospitals utilize iHOMIS, and more than a quarter utilize the UDRS. In addition, there are third-party providers, such as BizBox, MedSys, and Medix.

There are more health centers than hospitals that have a specific registry of diseases. Hospitals have better registries on CVD, stroke, heart attack, cancer, and mental disorders. However, health centers have better registries on hypertension, diabetes, respiratory tract diseases, and disability. More than half of the health centers visited utilize an online web-based database for patient records through iClinicSys, while only 1 uses another information system, specifically CHITS.

Most of the information systems utilized by the hospitals and health centers are for all patients in general wherein data on OPs can only be extracted. However, the extraction of data on OPs is complicated due to the nonuniform age group categories. In some facilities, the data on patients aged 60-64 years could not be properly retrieved, as these are incorporated into the 45-64-year age group. Age group classification is not standardized across facilities. Having multiple platforms for managing health information deteriorates interoperability between different health facilities, which, in effect, reduces the ease of service delivery.

Limitations

The study was able to cover facilities representing each region across the country; however, these are limited to the selected hospitals identified as geriatric centers and the health centers within their catchment area. Private health facilities were not covered by the study. Thus, the status of HISs in this study was limited only to public health facilities. Moreover, the status of the Philippines' HISs in general was not within the scope the study and thus warrants further investigation.

Comparison With Prior Work

In 1990, the BLACKBOX was the management information system for public health programs, vital statistics, mortality, and notifiable diseases. It handled and retrieved all data that were being routinely collected by public health workers all over the Philippines. It was developed toward a need-responsive and cost-effective health and management information system (HAMIS) [38]. Decades later, with the advancement of eHealth, there are various HISs in the country. For primary care benefit providers, there are 6 DOH-accredited EMR systems [36]. Furthermore, the DOH maintains 10 information systems and databases, which are being implemented in various health care settings through the UHMIS [31]. In addition, there are other private or third-party providers of HISs in the country. These are harmonized through the interoperability standards and guidelines issued by the DOH. However, based on the results of this study, there is no current database management and HIS specific for OPs to date.

The National Objectives for Health 2005-2010 and 2011-2016 prioritized the use of ICT in various reforms areas, critical health programs, and specific areas in health administration [39,40]. The Philippine eHealth Strategic Framework and Plan 2014-2020 was also developed [23]. The current and overall status of the PHIE warrants further investigation.

In terms of management of HISs, the most common reasons for not regularly updating the web-based database are a lack of or inadequate trained personnel to maintain and manage the information system, an unstable internet connection, the use of an offline system, and technical issues. These barriers are also consistent with the findings of other studies, such as a lack of standards, the use of different information systems, infrastructure issues for electricity and connectivity [35], a lack of human expertise [41], the need for training and support for human resources [41,42], and technical complexity [43,44]. In Malaysia, several issues have influenced overall HIS implementation in public hospitals, such as limited financial sources, maintenance by different departments, HIS implementation orders by the Malaysian Ministry of Health, addition of new systems, confidentiality issues, low acceptance levels, low satisfaction levels, different vendors, infrastructure issues, system breakdown, duplication of data, and different systems [45].

In developing countries, the establishment of well-coordinated information collection systems at various levels of the health care system using appropriate staff could contribute greatly to improvements in health care delivery [46]. Furthermore, ICT need to be seen as part of wider approaches involving

technological, social, and institutional innovation; health workers need to be educated more broadly on the use of HISs for action [46,47]; health institutions need to adapt in many ways toward local accountability and patient and health worker empowerment; and software development for HISs needs to integrate computerized systems with work practices to make work more effective [47]. In decentralized and democratic governments similar to the Philippines, HISs can play a crucial role in supporting and sustaining processes by serving as a repository for generated and analyzed information at the local level so that primary health care can address the dynamic and unpredictable elements of health care planning in developing countries [48].

Routine HIS interventions in the European region were identified to be promising; however, different areas of improvement, such as technical, organizational, and behavioral elements, were identified [49]. In Japan, the areas of improvement in health care information technology include the necessity for leadership and IT knowledge in medical communities, provider incentives, legislation regarding accountability, security, privacy and confidentiality, inclusion of stakeholders in solution development, and creation of sustainable business models [50].

In terms of sustainability of HISs, many challenges are faced, and these could be addressed through the systems' technical design, stakeholder coordination, and the building of organizational capacity to maintain and enhance such systems [51]. Furthermore, effective collaboration between major actors (donors, developers, and the Ministry of Health) is fundamental to sustain HISs [52].

Conclusion and Recommendations

The review of existing policies and guidelines provided a background on the status of HISs for OPs in the Philippines. The facility-based listing revealed the proportion of select facilities that conduct CGA and the status and challenges related to the HIS in select tertiary hospitals and health centers in the country.

Current HISs for OPs are characterized by fragmentation, multiple sources of health information, and inaccessibility. Barriers to achieving appropriate HISs for OPs include inability to update HISs in hospitals and health centers and a lack of age group and disease standardization.

A comprehensive assessment and care plan shared with all providers is one of the important elements of integrated care for OPs. In line with the universal health coverage and Sustainable Development Goal of "*Ensuring healthy lives and promote wellbeing for all at all ages*," an emerging landscape of innovation and development on integrated care of OPs is essential in order to address the multidimensional needs of the aging population.

A 1-person, 1-record EMR system for OPs is recommended in order to address their complex needs, as well as extract data to inform policies and programs. Furthermore, the data on OPs should be disaggregated and analyzed across geographic and social parameters in order to identify gaps in programs and provision of services.

Specifically, we recommend the following:

- Integration of data of OPs in the existing HISs in the country, wherein data can be derived and disaggregated across all health care facilities
  - Standardizing the definition of age groups (young, middle, and oldest old) and geriatric syndromes (ie, frailty, malnutrition, falls, dementia, delirium, incontinence, polypharmacy, deconditioning) and inclusion in the latest version of the NHDD (KMITS-DOH)
  - Funding and creating a dashboard for OPs (DOH, PhilHealth)
  - Conducting a CGA of all OPs as a clinical record to be shared across health care providers in all health settings, which will be integrated in the existing HIS
- Alignment of the integration of HISs for OPs with the existing mandates of the NHWPSC and health care provider networks (NHWPSC-DOH, centers for health development [CHDs], LGUs)
- Hiring and capacity building of personnel for management and maintenance of facility-based HISs (Health Human Resource Development Bureau [HHRDB]-DOH, regional hospitals, LGUs)
- Research, evaluation, and monitoring of the integrated HIS (National Commission of Senior Citizens [NCSC], National Privacy Commission [NPC], Health Policy Development and Planning Bureau [HPDPB]-DOH, academia, research institutions)

Acknowledgments

The authors thank the Advancing Health Through Evidence-Assisted Decisions (AHEAD) program of the Department of Health (DOH; specifically Dr Beverly Ho, Mr Pio Asuncion, and Ms Barbara Michelle de Guzman) and the Philippine Council for Health Research and Development, Department of Science and Technology (specifically Dr Jaime Montoya, Dr Paul Ernest De Leon, Ms Guiditta Galera, Ms Clarissa Reyes, and Mr Jay Vee Bartolome) for funding and fund management.

We also thank the National Institutes of Health, University of the Philippines Manila, for housing the implementing agency, which is the Institute on Aging, as well as the University of the Philippines (UP) for providing administrative assistance through the UP Manila channels and the Research Grants and Administration Office.

The Study 1 advisory board members and consultants were Dr Minerva Vinluan, Mr Roderick Napulan, Teresa Cleofe, Arlene Ruiz, Dennis Destacamento, Dr Orville Solon, Dr Nicamil Sanchez, Dr Harvy Joy Liwanag, and Atty. Maryann Lorraine Legaspi.

The authors also thank the hospital administrators of private hospitals, medical center chiefs and focal persons of the DOH regional hospitals, and mayors and city/municipal health officers and staff of health units that the team visited for their hospitality and active participation; policy makers, hospital administrators, and primary care implementers and older persons (OPs) and health care provider participants, barangay captains, and heads of the senior citizen organizations for their participation and kind assistance in the project implementation; and statisticians and project assistants (Leo Angelo Ocampo, Sarah Jane Fabito, Joy Marie Gillera, Maverick Tyrone Jacalan, Celia Abbago, Francis Michael Torreblanca [physical therapist registered in the Philippines (PTRP)], Lance Frederick Dusaban [registered pharmacist (RPh)], Xyzra Kyte Bober, Basil John Fitzgerald Zano, and Alyssa Mae Bernasol) and local field enumerators for their valuable assistance in data collection.

Finally, most of all, we thank the OPs for serving as the team's inspiration in this endeavor.

Authors' Contributions

APG performed protocol development, data collection, analysis, manuscript writing, and review and approval of the paper; SFDLV performed protocol development, data collection, manuscript writing, and review and approval of the paper; and SPM performed data collection, analysis, manuscript writing, and review and approval of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Ethics approval compilation.
[PDF File (Adobe PDF File), 24299 KB - jmir_v24i2e29541_app1.pdf]

Multimedia Appendix 2
Checklist for hospitals and health centers.
[PDF File (Adobe PDF File), 1053 KB - jmir_v24i2e29541_app2.pdf]

References

1. World Health Organization. Aging and Health: Key Facts. 2018 Feb 01. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2020-05-05]
2. Philippines Statistics Authority. Facts on Senior Citizens: Results of the 2015 Census of Population (Population Characteristics). 2017 Jun 01. URL: https://psa.gov.ph/system/files/2015%20Fact%20Sheets%20on%20Senior%20Citizen_pop.pdf?widt [accessed 2020-05-01]
3. World Health Organization. Strategic objective 5: Improving measurement, monitoring and research on healthy ageing. In: Global Strategy and Action Plan on Ageing and Health. Geneva: World Health Organization; 2017:20-23.
4. Ward K, Reuben D, Schmader K, Givens J. Comprehensive Geriatric Assessment. 2021 Jul 21. URL: https://www.uptodate.com/contents/comprehensive-geriatric-assessment [accessed 2021-03-01]
5. Stuck A, Siu A, Wieland G, Rubenstein L, Adams J. Comprehensive geriatric assessment: a meta-analysis of controlled trials. Lancet 1993 Oct 23;342(8878):1032-1036. [doi: 10.1016/0140-6736(93)92884-v] [Medline: 8105269]
6. Kuo H, Scandrett KG, Dave J, Mitchell SL. The influence of outpatient comprehensive geriatric assessment on survival: a meta-analysis. Arch Gerontol Geriatr 2004;39(3):245-254. [doi: 10.1016/j.archger.2004.03.009] [Medline: 15381343]
7. Stuck AE, Egger M, Hammer A, Minder CE, Beck JC. Home visits to prevent nursing home admission and functional decline in elderly people: systematic review and meta-regression analysis. JAMA 2002 Mar 27;287(8):1022-1028. [doi: 10.1001/jama.287.8.1022] [Medline: 11866651]
8. Elkan R, Kendrick D, Dewey M, Hewitt M, Robinson J, Blair M, et al. Effectiveness of home based support for older people: systematic review and meta-analysis. BMJ 2001 Oct 29;323(7315):719-725 [FREE Full text] [doi: 10.1136/bmj.323.7315.719] [Medline: 11576978]
9. Huss A, Stuck AE, Rubenstein LZ, Egger M, Clough-Gorr KM. Multidimensional preventive home visit programs for community-dwelling older adults: a systematic review and meta-analysis of randomized controlled trials. J Gerontol A Biol Sci Med Sci 2008 Mar;63(3):298-307. [doi: 10.1093/gerona/63.3.298] [Medline: 18375879]
10. Bachmann S, Finger C, Huss A, Egger M, Stuck AE, Clough-Gorr KM. Inpatient rehabilitation specifically designed for geriatric patients: systematic review and meta-analysis of randomised controlled trials. BMJ 2010 May 20;340:c1718 [FREE Full text] [doi: 10.1136/bmj.c1718] [Medline: 20406866]
11. Nord M, Lyth J, Alwin J, Marcusson J. Costs and effects of comprehensive geriatric assessment in primary care for older adults with high risk for hospitalisation. BMC Geriatr 2021 Apr 21;21(1):263 [FREE Full text] [doi: 10.1186/s12877-021-02166-1] [Medline: 33882862]
12. Department of Health. National eHealth Context (Overview). URL: http://ehealth.doh.gov.ph/#openModal9 [accessed 2020-11-10]

13. Dayrit M, Lagrada L. Philippines health system review 2018. In: Patcharanarumol W, Tangcharoensathien V, editors. Health Systems in Transition Vol. 8 No. 2. New Delhi: World Health Organization; 2018:1164.
14. Antonio C, Patdu I, Marcelo A. Health information privacy in the Philippines: trends and challenges in policy and practice. Acta Med Philipp 2016 Dec 31;50(4):223-236 [FREE Full text] [doi: 10.47895/amp.v50i4.760]
15. Vedel I, Akhlaghpour S, Vaghefi I, Bergman H, Lapointe L. Health information technologies in geriatrics and gerontology: a mixed systematic review. J Am Med Inform Assoc 2013 Nov;20(6):1109-1119 [FREE Full text] [doi: 10.1136/amiajnl-2013-001705] [Medline: 23666776]
16. Vest J, Kern L, Silver M, Kaushal R, HITEC investigators. The potential for community-based health information exchange systems to reduce hospital readmissions. J Am Med Inform Assoc 2015 Mar;22(2):435-442. [doi: 10.1136/amiajnl-2014-002760] [Medline: 25100447]
17. Vest JR, Unruh MA, Shapiro JS, Casalino LP. The associations between query-based and directed health information exchange with potentially avoidable use of health care services. Health Serv Res 2019 Oct;54(5):981-993 [FREE Full text] [doi: 10.1111/1475-6773.13169] [Medline: 31112303]
18. Department of Health. Philippine Health Facility Development Plan 2017-2022. URL: https://doh.gov.ph/sites/default/files/publications/HEALTH%20FACILITY%20DEVELOPMENT%20PLAN_0.pdf [accessed 2019-12-10]
19. Department of Health. Administrative Order No. 2017-0012: Guidelines on the Adoption of Baseline Primary Health Care Guarantees for All Filipinos. 2017 Aug 03. URL: https://dmas.doh.gov.ph:8083/Rest/GetFile?id=337131 [accessed 2019-06-16]
20. World Health Organization. World Report on Aging and Health 2015. 2015. URL: https://apps.who.int/iris/bitstream/handle/10665/186463/9789240694811_eng.pdf;jsessionid=AC48F87073D0B5F4FE80F53AB6FB488E?sequence=1 [accessed 2019-07-10]
21. World Health Organization. Monitoring the Building Blocks of Health System: A Handbook of Indicators and Their Measurement Strategies. Geneva, Switzerland: World Health Organization; 2010:1-110.
22. Republic of the Philippines. Universal Health Care Act: Republic Act No 11223. 2019 Feb 20. URL: https://www.officialgazette.gov.ph/downloads/2019/02feb/20190220-RA-11223-RRD.pdf [accessed 2019-03-10]
23. Department of Health. Philippine eHealth Strategic Framework and Plan 2014-2020. 2014 Apr 15. URL: http://ehealth.doh.gov.ph/index.php?option=com_content&view=category&layout=blog&id=73 [accessed 2020-11-09]
24. Department of Health. Joint DOH-DOST-PhilHealth Administrative Order No2016-001: Implementation of the Philippine Health Information Exchange. 2016 Jan 20. URL: http://ehealth.doh.gov.ph/index.php/poilicies/signed/69-aophie09102015 [accessed 2020-08-10]
25. Department of Health, Department of Science and Technology, Philippine Health Insurance Corporation. Joint Administrative Order No 2016-003: Adoption of the Philippine Health Information Exchange (PHIE) Lite. 2016 Jun 22. URL: http://ehealth.doh.gov.ph/index.php/poilicies/signed/72-1jointao20160003 [accessed 2020-08-10]
26. Department of Health. Administrative Order No 2013 -025: National Implementation of the Health Data Standards for eHealth Standardization and Interoperability (eHSI Release 001). 2013 Sep 10. URL: http://ehealth.doh.gov.ph/index.php/poilicies/signed/68-ao20130025nationalhealthdatastandards [accessed 2020-08-10]
27. Department of Health. Administrative Order No 2015-037: National Implementation of the Health Data Standards for eHealth Standardization and Interoperability. Manila, Philippines: Department of Health; 2015 Sep 02. URL: http://ehealth.doh.gov.ph/images/eHealthPDF/AO20150037.pdf [accessed 2020-07-10]
28. Department of Health. National Health Data Dictionary: Unified Health Management Information System. URL: https://uhmis.doh.gov.ph/index.php/downloads/home.php [accessed 2019-10-01]
29. Department of Health. DOH Programs: Health and Wellness Program for Senior Citizens (HWPSC). URL: https://www.doh.gov.ph/Health-and- Wellness-Program-for-Senior-Citizens [accessed 2019-10-07]
30. National Irrigation Administration, MIMAROPA. Overview: MIMAROPA Region. URL: http://region4b.nia.gov.ph/?q=content/overview [accessed 2022-01-31]
31. Department of Health. Unified Health Management Information System. URL: https://uhmis.doh.gov.ph/index.php/information-sys [accessed 2020-05-02]
32. BizBox Inc.. Why Bizbox?. URL: http://www.bizbox.ph/ [accessed 2020-01-15]
33. Medsys. Hospital Information Systems. URL: https://www.gomedsys.com/ [accessed 2020-01-15]
34. Medix. The Heart of Better Clinic Management. URL: https://www.medix.ph/learn-more [accessed 2020-01-15]
35. Ongkeko, Jr. AM, Fernandez RG, Sylim PG, Amoranto AJP, Ronquillo-Sy M, Santos ADF, et al. Community Health Information and Tracking System (CHITS): lessons from eight years implementation of a pioneer electronic medical record system in the Philippines. Acta Med Philipp 2016 Dec 31;50(4):264-279 [FREE Full text] [doi: 10.47895/amp.v50i4.769]
36. Philippine Health Insurance Corporation. PhilHealth Advisory No. 201600040: Electronic Medical Records System. 2016. URL: https://www.philhealth.gov.ph/advisories/2016/adv2016-0040.pdf [accessed 2020-06-20]
37. Department of Health. DOH Administrative Order No. 2017-0001: Policy Guidelines on the Standards of Care for Older Persons in All Healthcare settings. 2017 Jan 30. URL: https://dmas.doh.gov.ph:8083/Rest/GetFile?id=337116 [accessed 2020-05-26]

38. Marte BA, Schwefel D. The Philippine management information system for public health programs, vital statistics, mortality and notifiable diseases. Int J Biomed Comput 1995 Oct;40(2):107-114. [doi: 10.1016/0020-7101(95)01133-y] [Medline: 8847117]
39. Department of Health. National Objectives for Health: Philippines, 2005-2010. Manila, Philippines: Department of Health; Oct 2005:1-311.
40. Department of Health. Health Sector Reform Agenda Monograph No. 12. In: National Objectives for Health Philippines, 2011-2016. Manila, Philippines: Health Policy Development and Planning Bureau (HPDPB), Department of Health; Jul 01, 2012:1-149.
41. Williams F, Boren S. The role of the electronic medical record (EMR) in care delivery development in developing countries: a systematic review. Inform Prim Care 2008;16(2):139-145 [FREE Full text] [doi: 10.14236/jhi.v16i2.685] [Medline: 18713530]
42. Tomasi E, Facchini L, Maia M. Health information technology in primary health care in developing countries: a literature review. Bull World Health Organ 2004 Dec;82(11):867-874 [FREE Full text] [Medline: 15640923]
43. Ebardo R, Celis, NJ. Barriers to the adoption of electronic medical records in select Philippine hospitals. 2019 Apr 19 Presented at: Proceedings of the 5th International Conference on Computing and Artificial Intelligence; 2019; Bali, Indonesia p. 123-128 URL: https://dl.acm.org/doi/proceedings/10.1145/3330482 [doi: 10.1145/3330482.3330503]
44. Ismail NI, Abdullah NH, Shamsuddin A. Adoption of hospital information system (HIS) in Malaysian public hospitals. Procedia Soc Behav Sci 2015 Jan;172:336-343 [FREE Full text] [doi: 10.1016/j.sbspro.2015.01.373]
45. Ismail NI, Abdullah NH, Shamsudin A, Ariffin NAN. Implementation differences of hospital information system (HIS) in Malaysian public hospitals. IJSSH 2013;3(2):115-120 [FREE Full text] [doi: 10.7763/ijssh.2013.v3.208]
46. Azubuike M, Ehiri J. Health information systems in developing countries: benefits, problems, and prospects. J R Soc Promot Health 1999 Oct;119(3):180-184. [doi: 10.1177/146642409911900309] [Medline: 10518358]
47. Walsham G. Health information systems in developing countries: some reflections on information for action. Inf Technol Dev 2019 Mar 04;26(1):194-200 [FREE Full text] [doi: 10.1080/02681102.2019.1586632]
48. Madon S, Krishna S, Michael E. Health information systems, decentralisation and democratic accountability. Public Admin Dev 2010 Jun 08;30(4):247-260 [FREE Full text] [doi: 10.1002/pad.571]
49. Saigí-Rubió F, Pereyra-Rodríguez JJ, Torrent-Sellens J, Eguia H, Azzopardi-Muscat N, Novillo-Ortiz D. Routine health information systems in the European context: a systematic review of systematic reviews. Int J Environ Res Public Health 2021 Apr 27;18(9):1-26 [FREE Full text] [doi: 10.3390/ijerph18094622] [Medline: 33925384]
50. Abraham C, Nishihara E, Akiyama M. Transforming healthcare with information technology in Japan: a review of policy, people, and progress. Int J Med Inform 2011 Mar;80(3):157-170. [doi: 10.1016/j.ijmedinf.2011.01.002] [Medline: 21292546]
51. Moucheraud C, Schwitters A, Boudreaux C, Giles D, Kilmarx PH, Ntolo N, et al. Sustainability of health information systems: a three-country qualitative study in southern Africa. BMC Health Serv Res 2017 Jan 10;17(1):23 [FREE Full text] [doi: 10.1186/s12913-016-1971-8] [Medline: 28073361]
52. Kimaro H, Nhampossa J. The challenges of sustainability of health information systems in developing countries: comparative case studies of Mozambique and Tanzania. J Health Inform Dev Ctries 2007 Dec 18:1-10 [FREE Full text]

Abbreviations

AO: Administrative Order
CHITS: Community Health Information Tracking System.
CGA: comprehensive geriatric assessment
DILG: Department of Interior and Local Government
DOH: Department of Health
DOH EA: Department of Health Enterprise Architecture
EDPMS: Electronic Drug Price Monitoring System
EMR: electronic medical record
ePCB: expanded primary benefit care
FITforFrail: Focused Interventions for Frail Older Adults Research and Development Project
GSAP: Global Strategy and Action Plan on Aging and Health
HAMIS: health and management information system
HIE: health information exchange
HIS: health information system
iClinicSys: Integrated Clinic Information System.
ICNCDRS: Integrated Chronic Non-Communicable Disease Registry System
ICT: information and communication technologies
IDTOMIS: Integrated Drug Test Operations Management Information System
iHOMIS: Integrated Hospital Operations and Management Information System.
ITIS: Integrated TB Information System

KMITS: Knowledge Management and Information Technology Service
LGU: local government unit
NaRIS: National Rabies Information System
NHDD: National Health Data Dictionary
NHWPSC: National Health and Wellness Program for Senior Citizens
ONEISS: Online National Electronic Injury Surveillance System
OP: older person
PCHRD: Philippine Council for Health Research and Development
PHIE: Philippine Health Information Exchange
PHREB: Philippine Health Research Ethics Board
PRPWD: Philippine Registry for Persons with Disabilities
REC: research ethics committee
RHU: rural health unit
Seg-RHIS: Segworks Tecknologies
SHINE OS+: Secure Health Information Network and Exchange
SJREB: Single Joint Research Ethics Board
UDRS: Unified Disease Registry System.
UHC: universal health care
UHMIS: Unified Health Management Information System
UPMREB: University of the Philippines Manila Research Ethics Board
VAWCRS: Violence Against Women and Children Registry System
WAH: Wireless Access for Health
WHO: World Health Organization

Edited by R Kukafka; submitted 12.04.21; peer-reviewed by D Emwodew, F Sukums; comments to author 15.05.21; revised version received 17.06.21; accepted 05.07.21; published 14.02.22.

Please cite as:
Garcia AP, De La Vega SF, Mercado SP
Health Information Systems for Older Persons in Select Government Tertiary Hospitals and Health Centers in the Philippines: Cross-sectional Study
J Med Internet Res 2022;24(2):e29541
URL: https://www.jmir.org/2022/2/e29541
doi:10.2196/29541
PMID:35156927

Original Paper

Habit and Automaticity in Medical Alert Override: Cohort Study

Le Wang[1], PhD; Kim Huat Goh[2], PhD; Adrian Yeow[3], PhD; Hermione Poh[4], BSc (Hons); Ke Li[4], MD; Joannas Jie Lin Yeow[4], BBus; Gamaliel Tan[4,5], MMed, MBBS; Christina Soh[2], PhD

[1]City University of Hong Kong, Hong Kong, China (Hong Kong)
[2]Nanyang Technological University, Singapore, Singapore
[3]Singapore University of Social Sciences, Singapore, Singapore
[4]Medical Informatics, National University Health System, Singapore, Singapore
[5]Ng Teng Fong General Hospital, Singapore, Singapore

Corresponding Author:
Kim Huat Goh, PhD
Nanyang Technological University
50 Nanyang Avenue
Singapore, 639798
Singapore
Phone: 65 67904808
Email: akhgoh@ntu.edu.sg

Abstract

Background: Prior literature suggests that alert dismissal could be linked to physicians' habits and automaticity. The evidence for this perspective has been mainly observational data. This study uses log data from an electronic medical records system to empirically validate this perspective.

Objective: We seek to quantify the association between habit and alert dismissal in physicians.

Methods: We conducted a retrospective analysis using the log data comprising 66,049 alerts generated from hospitalized patients in a hospital from March 2017 to December 2018. We analyzed 1152 physicians exposed to a specific clinical support alert triggered in a hospital's electronic medical record system to estimate the extent to which the physicians' habit strength, which had been developed from habitual learning, impacted their propensity toward alert dismissal. We further examined the association between a physician's habit strength and their subsequent incidences of alert dismissal. Additionally, we recorded the time taken by the physician to respond to the alert and collected data on other clinical and environmental factors related to the alerts as covariates for the analysis.

Results: We found that a physician's prior dismissal of alerts leads to their increased habit strength to dismiss alerts. Furthermore, a physician's habit strength to dismiss alerts was found to be positively associated with incidences of subsequent alert dismissals after their initial alert dismissal. Alert dismissal due to habitual learning was also found to be pervasive across all physician ranks, from junior interns to senior attending specialists. Further, the dismissal of alerts had been observed to typically occur after a very short processing time. Our study found that 72.5% of alerts were dismissed in under 3 seconds after the alert appeared, and 13.2% of all alerts were dismissed in under 1 second after the alert appeared. We found empirical support that habitual dismissal is one of the key factors associated with alert dismissal. We also found that habitual dismissal of alerts is self-reinforcing, which suggests significant challenges in disrupting or changing alert dismissal habits once they are formed.

Conclusions: Habitual tendencies are associated with the dismissal of alerts. This relationship is pervasive across all levels of physician rank and experience, and the effect is self-reinforcing.

(J Med Internet Res 2022;24(2):e23355) doi:10.2196/23355

KEYWORDS

alert systems; habits; electronic medical record; health personnel alert fatigue

Introduction

Background and Significance

Electronic medical records (EMR) systems have many embedded clinical support alerts to help warn or remind clinicians about patient-related issues [1]. However, the ubiquitous use of such alerts has led to a significant number of alert dismissals [2-8]. Some of these alerts were wrongly dismissed, leading to significant health consequences [1,9]. For example, Slight et al [10] estimated that in the United States alone about 57.6 million adverse drug event alerts were dismissed in 2014, and of those, about 5.5 million were inappropriately overridden, resulting in approximately 196,600 adverse drug events.

Prior literature suggests at least 3 interrelated reasons for the excessive dismissal of alerts. The first reason is the relevance or effectiveness of the alerts [11,12]. Physicians are more likely to dismiss less relevant alerts when they are repeatedly exposed to them. As such, studies have examined how the tiering of alerts based on their medical significance [13] could improve a physician's alert compliance. Irrelevant alerts are also linked to the second reason—alert fatigue. Alert fatigue is a result of alert overload that causes clinicians to become desensitized to subsequent alerts [14,15]. The third reason, as suggested by Baysari et al [16], alerts are excessively dismissed is because of physicians' habitual dismissal of alerts. Using field observations and physician interview data, the authors found that physicians had developed the habit of dismissing alerts over time, which resulted in an excessive number of alerts being dismissed without significant cognitive considerations. As such, Baysari et al [16] called for more empirical studies to examine the role that habit plays in influencing alerts dismissal and establish the prevalence of such incidences of habitual dismissal. In particular, they suggested studying the association between the number of alerts clinicians experience and their alert dismissal rates and how the rate of alert exposure impacts the formation of alert dismissal habits.

Put together, if physicians under high workload environments rely on habituated responses as a way to cope with alert overload and not critically process these alerts as a result [16], then it would have significant implications on the efficacy of alerts in situations that matter most. Therefore, we concur with Baysari et al [16] that it is important to establish the prevalence of such habitual behaviors among physicians and understand how such habits are formed and their impacts on patient care.

Habits are driven by an environmental stimulus that leads to consistent follow-up action in response to that stimulus. Quantifying habitual behavior using empirical observations is challenging; however, this area of research has received increased attention in recent years with the development of quantitative models to measure habit strength [17,18]. These models permit researchers to empirically quantify habitual actions to answer the call for empirically investigating the effects of habits on alert dismissal [16].

Objective

The objective of this study is to empirically quantify the prevalence of alert dismissal associated with physician habit (habitual dismissal) as well as the association between a physician's habitual dismissal and their subsequent tendency for alert dismissal.

Methods

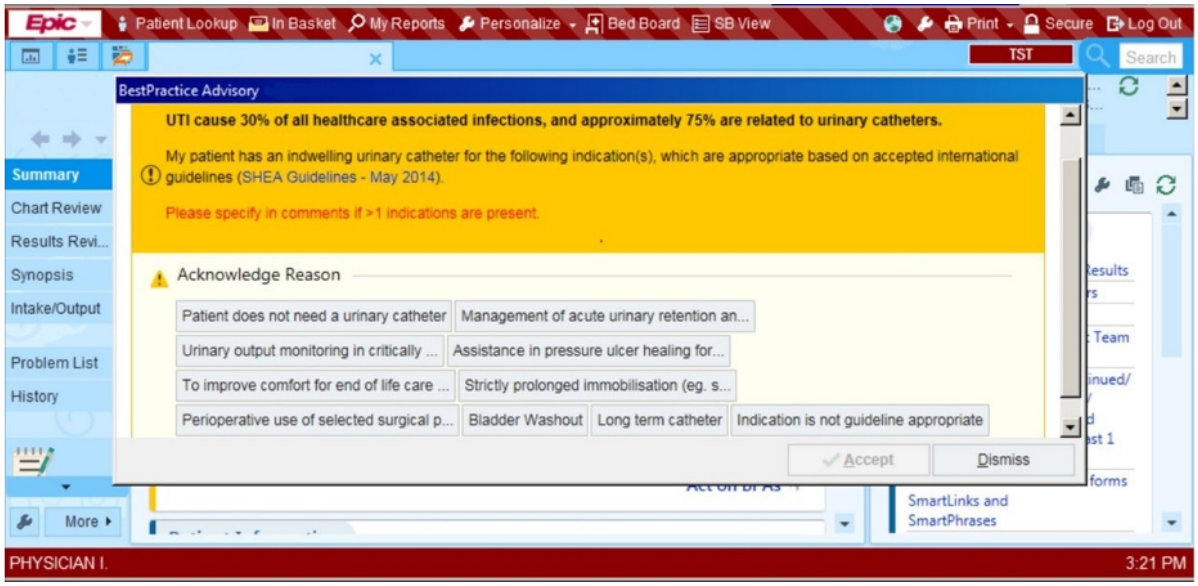
Research Context

The study is a retrospective analysis of log data from 66,049 alerts generated from hospitalized patients captured by the Epic EMR system (Epic Systems Corp) over a 20-month period from March 2017 to December 2018.

Settings and Data Context

The study was situated at Jurong Health Campus, which consists of the integrated 700-bed Ng Teng Fong General Hospital and the 400-bed Jurong Community Hospital. An alert was set up in the EMR system to check if a patient had an indwelling catheter (IDC) for an extended period and was at risk of urinary tract infection. The IDC alert was activated each time the physician accessed the patient's medical record, and it would have informed the physician of the following clinical guideline: "Urinary tract infection causes 30% of all health care–associated infections, and approximately 75% are related to urinary catheters." A screenshot of the alert can be seen in Figure 1. The alert would have reminded the physician to assess the patient and indicate the reasons as to whether the patient should or should not continue to use the catheter. If the physician indicated that the catheter should be removed, they could enter the order in the Epic system. If the patient required continued use of the catheter, the physician could acknowledge this and indicate the reasons in the alert pop-up. In order to not obstruct the physician's workflow, the alert also allowed the physician to ignore the alert by clicking on the dismiss button. Alerts that were acted upon by physicians (eg, alerts that were acknowledged with reasons for nonremoval or ordered for removal) would no longer appear. Alerts dismissed without acknowledgment continued to appear each time the physician accessed the patient's medical record.

Figure 1. Indwelling catheter alert interface screenshot (modified to remove patient identifiers).



Between March 2017 and December 2018, we captured the actions of 1152 physicians who interacted with the IDC alert. This IDC alert was implemented in March 2017, after which we began the data capture. Physicians were not exposed to this alert prior to March 2017 as the IDC alert was custom-built and implemented only in this hospital campus. The IDC alert's interface was also different from other clinical alerts in the EMR system as it is a customized alert. Our data consist of a total of 66,049 alerts from 1874 patients; each patient triggered an average of more than 30 alerts during their hospitalization. A total of 91.2% (60,236/66,049) of all alerts were dismissed by physicians.

Outcome: Alert Dismissal

The unit of analysis for this study is each alert instance, and the main dependent variable is alert dismissal (*dismiss*), represented by a binary measure to indicate if the physician dismisses the alert. As a robustness check, we have also included 3 other possible dependent variables: dismissal performed in under 1, 2, or 3 seconds after the alert's appearance (*dismiss1*, *dismiss2*, and *dismiss3*). We have included time limits for these alerts as a robustness check because habits are often associated with actions that are performed automatically with little cognitive effort and awareness.

Automaticity [16,19] is a trait of habits and an efficient way of handling familiar situations where the individual can execute actions with limited mental effort. Automaticity during a habitual action means that an individual is less likely to search for new information while ignoring additional situational information and will follow up with the usual habitual responses in a short amount of time [20]. Prior literature suggests that habitual dismissal occurs when physicians dismiss an alert without reading the content of the alert [16,21]. Given the clinical information included in the IDC alert, dismissal of alerts in under 1 second after they appeared were likely to be acts of automaticity and it was unlikely the physician had assessed the content of the alert. As a robustness check, other timing cutoffs were also used (ie, alerts dismissed in under 2 seconds [*dismiss2*] and 3 seconds [*dismiss3*]) with statistically similar findings.

Primary Predictors: Habit Strength From Habitual Learning

To quantify habitual behavior, we represent habitual learning as a mathematical form of Hebbian learning as proposed in prior studies [17,22-26].

$$H_{t+1} = H_t + \alpha_H (\alpha_t - H_t) \quad \textbf{(1)}$$

H represents the physician's habit strength at different points in time denoted by t. α_t corresponds to the action taken on time t. An action that corresponded to the response observed in the habitual behavior (ie, dismissal of alert) was coded as 1; it was coded as 0 if the action was otherwise. α_H is a parameter that quantifies the rate of learning after each habitual action. Larger α_H values suggest a faster habit learning rate, and it typically takes on small values (ie, less than 1) [17]. For robustness and comparative purposes, this study used 3 values of α_H to compute the physician's habit strength—0.01 ($H_{0.01}$), 0.05 ($H_{0.05}$), and 0.1 ($H_{0.1}$). Given that the alert system is custom-built and that we captured all alert responses from the start of the alert's implementation, we set the initial habit strength, H_t, at 0 since no physicians had prior exposure to this particular alert before the implementation of the alert.

Covariates

We divided the covariates used in this study into 5 categories: context of the alert, physician's historical exposure to alerts, physician's characteristics, patient's characteristics, and timing effects. The names of the covariates are in brackets and italicized.

- Context of the alert: captured if the alert appeared during ward rounds (ward).
- Physician's historical exposure to alerts were controlled for in 5 ways. First, given that many physicians can consult 1 patient and a physician can have many patients, we first computed the patient-physician (P_C) dyads. This dyad allowed us to precisely measure each physician's alert exposure for a particular patient. We computed P_Ctotal,

the total number of alerts a particular physician received for a particular patient. Second, we computed P_Cward, the total number of alerts received during ward rounds for the patient-physician pair. Third, to control for prior exposure to alerts unrelated to a particular patient, we computed the total number of alerts a particular physician received within the hospital from the beginning of the IDC alert program's implementation (Ctotal). Physicians have heterogeneous experience and exposure to these alerts as they consult a large number of patients; the Ctotal variable controls for a physician's familiarity with the system. Fourth, we controlled for the number of unique patients with indwelling catheters the particular physician encountered in a particular workday (CPatNum) to account for the physician's exposure to patients with similar ailments. Finally, we controlled for the number of days since the physician,C, received the first alert for the patient, P.

- Physician characteristics: the variable C_rank (ie, seniority of the physician—intern, resident, fellow, or attending) was a proxy for the level of physician experience. We also captured the main medical specialty (specialty) in charge of the patient (eg, cardiology). Finally, we captured the type of ward the patient was in (dept).
- Patient characteristics: the patient's age, gender, and race; the total length of stay (los) and number of diagnoses based on International Classification of Diseases, Tenth Revision classification (Diagnosis_count) were used as proxies for the severity of the patient's condition.
- Timing effects: controlled for by recording the day of the week to account for the change in shift duty and the month of the year to account for seasonality.

The complete list of covariates and their detailed descriptions can be seen in Table S1 in Multimedia Appendix 1.

Statistical Analysis

We computed the descriptive statistics to explore the distribution of how long it took each alert to be processed across different physician types and alert outcomes. We computed the point biserial correlations between the outcome variables and primary predictors, as well as Pearson correlations among the rest of the primary predictors. Calculating the point biserial correlation was required for instances where one of the variables was a binary variable.

To test the independent association of the physician's habit strength on dismissal, we estimated 3 fixed effects, logistic regressions with the variable dismiss as the outcome variable, and each of the 3 measures of habit strength ($H_{0.05}$, $H_{0.01}$, and $H_{0.1}$) as predictor variables. Fixed effects regression—grouped at the physician level—was required as the same physician's dismissal behavior is likely to be correlated across alert instances, and controlling for physician-level effects allowed us to isolate the association between habit strength and dismissal outcomes beyond the physician's idiosyncratic characteristics. As a robustness test, we estimated 3 more sets of fixed effects: logistic regression models using alternative, dependent variable measurements of dismissal (ie, *dismiss1*, *dismiss2*, and *dismiss3*). Likewise, the 3 primary predictors of habitual learning were regressed on each of these 3 alternative dependent variables, resulting in an additional (3 × 3) 9 regressions. As a further robustness check, we replicated the 12 regression models described above using regular logistic estimators and random effects logistic estimators for comparative purposes. These additional 24 models are reported in Multimedia Appendix 1.

All regressions used the covariates described above as controls during estimation, and we performed our analyses using Stata (version 14.2, StataCorp LLC).

Ethics Approval

Ethics approval for this study was received by the Domain Specific Review Board (Ref: 2018/01306) in the National Healthcare Group, Singapore.

Results

Descriptive Statistics

Table 1 presents the summary statistics of our sample. We observed that physicians, on average, dismissed 91.2% (60,236/66,049) of all IDC alerts they encountered. Further, 13.3% (8750/66,049) of all alerts they encountered were dismissed in under 1 second, 56.9% (37,546/66,049) of alerts were dismissed in under 2 seconds, and 72.5% (47,890/66,049) of alerts were dismissed in under 3 seconds. A total of 45.6% (30,132/66,049) of all alerts appeared during the physician's ward rounds (*ward*). Physicians, on average, were exposed to 84.9 alerts during this period in the hospital (C_{total}). For any specific patient, a physician received, on average, about 3.9 alerts within the ward (P_C_{ward}) and 8.2 alerts during the patient's stay (P_C_{total}). On any particular day, a physician would encounter an average of 1.3 patients (C_{PatNum}) when the IDC alert was triggered.

Figure 2 presents the distribution of processing time for the alerts before they were dismissed. We observed that, on average, 79.5% (47,890/60,236) of all dismissed alerts were dismissed in under 3 seconds, and 14.5% (8750/60,236) of all dismissed alerts were dismissed in under 1 second.

We also observed changes in physicians' response times over time (Figure 3). The average time for a physician to process an alert dropped from 5.90 seconds (95% CI 5.46-6.34) in the first alert exposure to 2.43 seconds (95% CI 1.75-3.10) during the 60th exposure.

We further explored the distribution of dismissal times across different physician experience levels (Figure 4) and observed similar patterns in dismissal times across medical interns, residents, fellows, and attending physicians. These results suggest that habitual tendencies in dismissing alerts apply to all levels of medical experience.

To visualize how a physician's habit strength ($H_{0.01}$, $H_{0.05}$, and $H_{0.10}$) changed relative to their increasing exposure to the alerts, see the relationship plotted for a typical physician in Figure 5.

Table 1. Summary statistics of the variables[a].

Variable	Mean (SD)	Minimum	Maximum
dismiss	0.912 (0.283)	0	1
dismiss1	0.132 (0.339)	0	1
dismiss2	0.568 (0.495)	0	1
dismiss3	0.725 (0.446)	0	1
$H_{0.01}$	0.400 (0.292)	0	1
$H_{0.05}$	0.708 (0.316)	0	1
$H_{0.01}$	0.792 (0.291)	0	1
ward	0.456 (0.498)	0	1
P_C_{total}	8.216 (11.961)	1	120
P_C_{ward}	3.876 (5.991)	0	60
C_{total}	84.868 (93.886)	1	669
C_{PatNum}	1.303 (0.619)	1	7
day_lag	3.164 (8.826)	0	240
age	67.903 (16.125)	16	105
gender	0.558 (0.497)	0	1
los	38.425 (48.369)	1	357
Diagnosis_count	1.044 (0.234)	1	3

[a]All variables listed above were described earlier in the text. Refer to Table S1 in Multimedia Appendix 2 for a list of detailed definitions of the variables.

Figure 2. Distribution of processing time of all dismissed alerts.

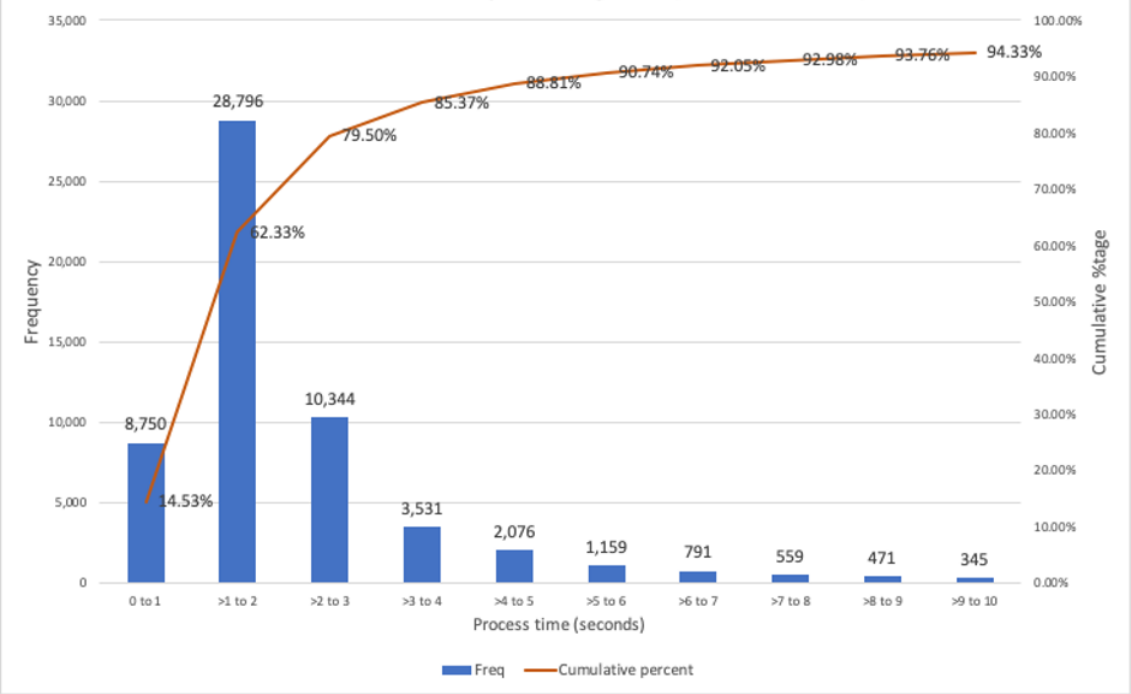


Figure 3. Physicians' average alert processing times across number of alert exposures.

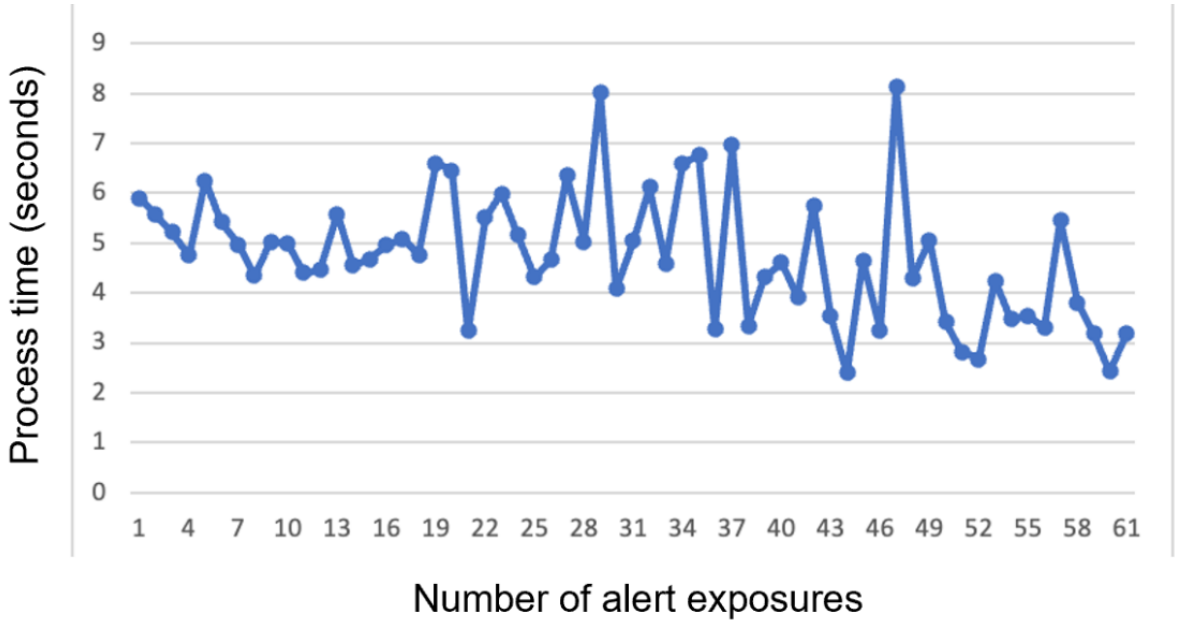


Figure 4. Distribution of response time by physician's rank (level of experience).

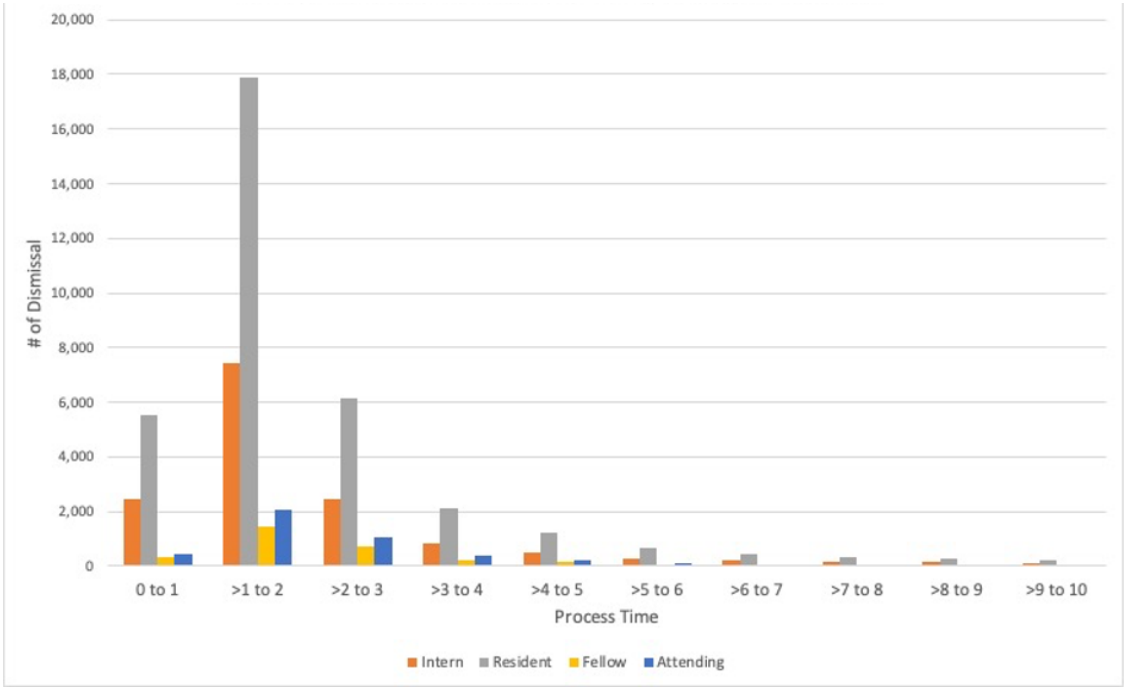
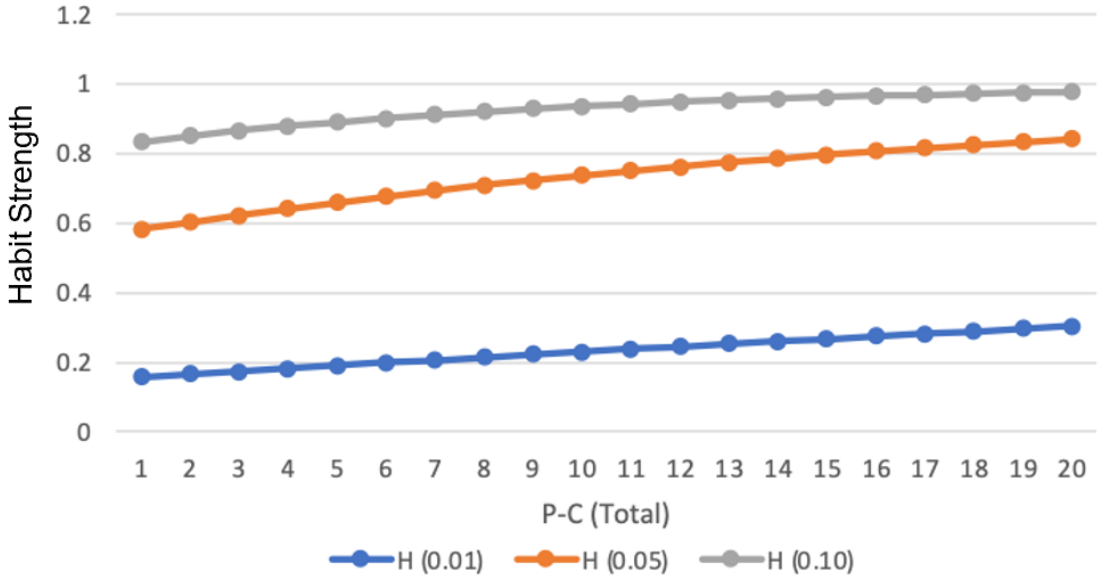


Figure 5. Variation of habit strength across number of alert exposures. Note: P-C (Total) represents the number of alert exposures a physician experiences.



The pair-wise correlations between the outcome variable and primary predictors can be found in Table 2. Alert dismissal was found to positively correlate with all primary predictors (habit strength) in this study, with correlations ranging from 0.227 to 0.421 (P<.001). Further, all dismissals under 1, 2, and 3 seconds (*dismiss#*) were found to positively correlate with the physicians' habit strengths. The point biserial correlation values ranged from 0.024 (P<.001) to 0.270 (P<.001). The correlations were found to be higher for dismissals under 2 seconds (the correlation values ranged from 0.106 to 0.220; P<.001) than dismissals under 1 second (the correlation values ranged from 0.024 to 0.082; P<.001). Dismissals under 3 seconds had the highest correlations between the physician's habit strength and alert dismissal (correlations ranged from 0.136 to 0.270; P<.001). We observed stronger positive correlations for a higher rate of habitual learning with dismissal outcomes, indicating that physicians with stronger habitual tendencies were more likely to dismiss alerts within a short period of time.

Table 2. Correlations[a] among outcome variables and key predictors.

	dismiss (P value)	*dismiss1* (P value)	*dismiss2* (P value)	*dismiss3* (P value)	$H_{0.01}$ (P value)	$H_{0.05}$ (P value)	$H_{0.1}$ (P value)
dismiss	.99	—[b]	—	—	—	—	—
dismiss1	0.121 (<.001)	.99	—	—	—	—	—
dismiss2	0.357 (<.001)	0.341 (<.001)	.99	—	—	—	—
dismiss3	0.505 (<.001)	0.241 (<.001)	0.707 (<.001)	.99	—	—	—
$H_{0.01}$	0.227 (<.001)	0.024 (<.001)	0.106 (<.001)	0.136 (<.001)	.99	—	—
$H_{0.05}$	0.347 (<.001)	0.065 (<.001)	0.183 (<.001)	0.223 (<.001)	0.845 (<.001)	.99	—
$H_{0.01}$	0.421 (<.001)	0.082 (.002)	0.220 (<.001)	0.270 (<.001)	0.703 (<.001)	0.956 (<.001)	.99

[a]Variables *dismiss*, *dismiss1*, *dismiss2*, and *dismiss3* are binary, so we perform a point biserial correlation for their relationships with other predictors. All other correlations are Pearson correlations.

[b]Not applicable.

Regression Results

The regression results in Multimedia Appendix 2 show the association between physicians' current habit strength of alert dismissal ($H_{0.01}$, $H_{0.05}$, $H_{0.1}$) and the probability of their dismissal of the current alert (*dismiss*). The results show that all 3 ways of quantifying habit strengths were positively associated with an increased likelihood of dismissing alerts received in the present time. A 1 standard deviation increase in habit strength was associated with an increase in odds of a physician's dismissal of an alert they received in the present time by 0.642 to 0.810 times.

The results for the other 9 models with alternative measures of dismissals performed within 1, 2, and 3 seconds of appearance (*dismiss1*, *dismiss2*, and *dismiss3*) were also consistent with the main findings. Physicians with higher habit strength were found to more likely dismiss the next alert in under 1 to 3 seconds. Specifically, every 1 standard deviation increase in habit strength was associated with an increase in odds ratio by 0.362 to 0.510 times for the dismissal of the next alert within 1 second of the

alert's appearance. The effect was similar for dismissals that occur in under 3 seconds of the alert's appearance—a 1 standard deviation increase in a physician's habit strength was associated with an increase in odds ratio of 0.350 to 0.503 times for dismissing the subsequent alert. We further computed a random-effects estimator to derive the intraclass correlation coefficient for this model and found that physician characteristics account for 0.716 of the total unexplained variance in dismissal.

Extant studies have found mixed results on how physicians of different ranks (ie, level of experience) would respond to alerts differently. For example, Baysari et al [27] found that only 17% of alerts presented to senior doctors were read compared to junior doctors reading 78% of patient alerts they received. However, Straichman et al [28] and Tamblyn et al [29] found that a physician's seniority had no impact on their alert overriding behavior. To explore if habits had a different influence on physicians of different ranks (levels of experience), we divided the alert instances into 4 subsamples based on the physician's rank (intern, resident, fellow, or attending). For each subsample, we observed that an increase in habit strength was positively associated with dismissals, with dismissals below 1 to 3 seconds of alert exposure across all physician ranks (Table 3).

Table 3. Fixed effects logistic regression results for different physician ranks.

Physician rank and habit strength	*dismiss*		*dismiss1*		*dismiss2*		*dismiss3*	
	β[a] (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
Intern (first year post-medical school)								
$H_{0.05}$	3.235 (2.785-3.685)	.23	1.656 (1.368-1.944)	.15	1.476 (1.288-1.664)	.10	1.560 (1.364-1.756)	.10
Resident (2 to 6 years of work experience)								
$H_{0.05}$	2.604 (2.374-2.835)	.12	1.354 (1.174-1.535)	.10	1.496 (1.374-1.618)	.06	1.486 (1.357-1.614)	.07
Fellow (completed residency and in specialist training)								
$H_{0.05}$	1.591 (0.372-2.811)	.62	2.285 (1.397-3.173)	.45	1.831 (1.333-2.329)	.25	1.584 (1.010-2.157)	.29
Attending physician (specialists)								
$H_{0.05}$	1.830 (1.008-2.651)	.42	1.934 (1.272-2.596)	.34	2.077 (1.641-2.512)	.22	1.652 (1.184-2.121)	.24

[a]Coefficients are exponentiated and represent odds ratios. Each cell represents the coefficient of a single fixed effects logistics regression. Full regression results are available in Multimedia Appendix 1. All multivariate models are adjusted for the context of the alert, physician's historical exposure to alerts, physician characteristics, patient characteristics, and timing effects.

Finally, we computed the relationship between habit strength and physician processing time in an alternative analysis with fixed effects, random effects, and ordinary least square estimators. Here, we observed that processing time of alert is significant and negatively associated with habit strength after controlling for contextual factors, providing further evidence for the association between habit and automaticity ($\beta_{fixed\_effects}$=–1.161, P<.10; $\beta_{random\_effects}$=–1.305, P<.05; β_{OLS}=–2.402, P<.05). We further computed a random-effects estimator to derive the intraclass correlation coefficient for this model and found that physician characteristics account for 0.164 of the total unexplained variance in processing time.

Discussion

Principal Findings

Physicians in clinical settings often experience high workload, significant time pressure, and information overload. The use of clinical support alerts supposedly ensures that, amid these challenging work conditions, physicians can quickly attend to patients potentially at risk of adverse events. The irony, however, is that the very same challenging conditions are also the conditions that may result in the formation of habits in physician responses to the alerts [16,20]. Under such conditions of high workload and information overload, individuals are more likely to rely on automaticity in their responses. Further, as physicians experience more alerts, as argued by Baysari et al [16], they may form habits that are biased toward more habitual dismissal. We empirically show that this is true in our study's clinical context.

Our study contributes to the area of habit research by operationalizing an empirical measure of habitual actions. Traditionally, habitual actions are observed and understood from the perspective of the physician performing the action. However, it is inherently difficult to determine the cognitive state of the physician as they perform different actions [19]. As such, it is challenging to conduct a large-scale empirical study to quantify physician habits. Our study provides a different approach to address this challenge by using a theoretically based alternative that tracks repeated actions as an empirical proxy to quantify habitual actions [18]. We tested the habit strength measure's relationship with the physician's subsequent alert dismissal and found significant positive relationships between them. Our additional analyses also showed that habit strength

measures yielded consistent and similar results with models that used simpler measurements of habit such as physicians' past dismissal rates. Our empirical method of quantifying the habitual response of alerts provides a way for future empirical research that studies the role of habits in physician actions in general.

Our study also answers the call for more systematic empirical research to examine the role of habits in high alert dismissal. Our findings, where we observed a large proportion of alerts dismissed in under 3 seconds, provide empirical evidence of physicians relying on automaticity while dismissing alerts. Our findings support the view that habits play a significant role in influencing alert response.

Specifically, our findings provide 3 key insights into the relationship between habits and alerts. First, we found that the association between habit strength and alert dismissal is pervasive in the sample of alert responses we studied. Second, we showed that habitual alert dismissal occurs in physicians across all levels of experience; senior and junior physicians alike all have the tendency to habitually dismiss alerts. Together, these insights suggest that health informatics professionals and designers of clinical support alerts need to take into account the effect of habits on physicians when designing, implementing, and interpreting the impact of clinical alerts. Research on alert relevance that uses dismissal rates as a measure of relevance (ie, lower dismissal rates suggest a higher relevance of the alert) need to take into account how fast alerts are dismissed. Our study found that there is more to alert dismissal than simply a physician's judgment of the alert's relevance. This is especially true for alerts dismissed within time durations that are too brief to support any meaningful assessment of the alert. Third, we found that habits are self-perpetuating, self-reinforcing, and, as a result, hard to mitigate once they are formed. Physicians who have a history of dismissing alerts are more likely to dismiss subsequent alerts and tend to dismiss them more hastily without much consideration. The self-reinforcing nature of habitual dismissal is a cause of concern as it may significantly reduce the effectiveness of the focal alert and possibly impact the physicians' responses to other clinical support alerts implemented in a clinical setting.

However, given our understanding of the role of a physician's habit formation in the context of alert dismissal and the prevalence of habitual dismissal in health care settings, we propose that future research should draw on our understanding to examine how work conditions and alerts could be designed to mitigate the formation of habitual behavior. Extant literature informs us that habits are shaped by associating an external cue with a repeated, stable response, which, in turn, produces a set of consistent consequences [30]. Habits form when individuals experience a stable context that triggers a habitual action performed frequently, and the habitual action leads to an affirmative outcome that reinforces the behavior. These 3 antecedents are necessary for habit formation, and prior studies on strategies to reduce habitual responses have suggested disrupting or removing these antecedents to reduce habitual behaviors [31,32]. One example to mitigate habit formation is to disrupt the stable environment and triggers that lead to habitual responses. Although these studies are not situated in the medical context, the strategy of removing the individual's familiar environment and triggers of habitual response is an approach that could be explored in the clinical alert context. For example, designers of clinical support alerts could consider varying the format and form of the alerts based on the risk levels or types of clinical conditions involved. Rare alerts associated with higher risks could be designed with a different alert interface, which could disrupt the stable environment (similar presentation mode) that triggers the habitual response. Furthermore, the alert interface could be programmed to be refreshed periodically to remove the stable environment required for habit formation and thereby reduce the incidence of habitual responses.

Limitations

One limitation of this study is that it examined habitual dismissal for only one type of clinical alert: the removal of IDCs. Although we examined all the alerts that appeared during the sample time frame, future research could include different types of alerts as part of the study to understand if our findings are generalizable beyond a specific alert. However, studies involving different alerts will require these additional alert systems to be custom-built to accurately analyze physician responses from the start of the alert's implementation to quantify physician habit strength. Studies that use existing alert systems may not yield accurate findings, as physicians would have had prior exposure to these alerts. Although we attempted to control for physicians' alert fatigue by including, as a covariate, the number of unique patients with the IDC alert the physician consulted on that day, this may not completely control for alert fatigue that might result from other types of medical alerts. Again, future research may attempt to include the impact of other alerts as a control in the study.

Finally, like most retrospective cohort studies, our study does not seek to establish causal evidence or a causal relationship between a physician's habits and the subsequent dismissal of alerts. Here, we use a computational model of habit to show the association between habit formation and a physician's dismissal of alerts within a real-world clinical context. Without further studies that examine the intentions, less feasible in a real-world clinical context, the ability to isolate habit as the sole cause of dismissal is challenging and calls for further research.

Conclusion

This study shows that the strength of a physician's habit for dismissing medical alerts is positively associated with their tendency for subsequent incidences of alert dismissal. Additionally, it was found that most (72.5%) dismissed alerts occurred in under 3 seconds of the physician's exposure to the alert. This empirical finding is in line with prior health informatics literature, suggesting the role of habit in EMR alert dismissal. We contribute to this stream of work by showing that habitual dismissal occurs across all levels of physician experience and that this is a self-reinforcing phenomenon. As physicians habitually dismiss alerts, the likelihood of them hastily dismissing subsequent alerts increases significantly. This phenomenon presents challenges to removing such inclinations toward habitual dismissal among physicians.

Acknowledgments

This research received funding from the Social Science Research Council, Ministry of Education, Singapore.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables.
[DOCX File , 137 KB - jmir_v24i2e23355_app1.docx]

Multimedia Appendix 2
Fixed effects logistics regression results for dismissal.
[DOCX File , 16 KB - jmir_v24i2e23355_app2.docx]

References

1. Russ A, Zillich A, Melton B, Russell S, Chen S, Spina J, et al. Applying human factors principles to alert design increases efficiency and reduces prescribing errors in a scenario-based simulation. J Am Med Inform Assoc 2014 Oct;21(e2):e287-e296 [FREE Full text] [doi: 10.1136/amiajnl-2013-002045] [Medline: 24668841]
2. van der Sijs H, Aarts J, Vulto A, Berg M. Overriding of drug safety alerts in computerized physician order entry. J Am Med Inform Assoc 2006 Mar 01;13(2):138-147. [doi: 10.1197/jamia.m1809]
3. Nanji KC, Slight SP, Seger DL, Cho I, Fiskio JM, Redden LM, et al. Overrides of medication-related clinical decision support alerts in outpatients. J Am Med Inform Assoc 2014 May 01;21(3):487-491 [FREE Full text] [doi: 10.1136/amiajnl-2013-001813] [Medline: 24166725]
4. Bryant A, Fletcher G, Payne T. Drug interaction alert override rates in the Meaningful Use era. Appl Clin Inform 2017 Dec 19;05(03):802-813. [doi: 10.4338/aci-2013-12-ra-0103]
5. Phansalkar S, van der Sijs H, Tucker A, Desai A, Bell D, Teich J, et al. Drug-drug interactions that should be non-interruptive in order to reduce alert fatigue in electronic health records. J Am Med Inform Assoc 2013 May 01;20(3):489-493 [FREE Full text] [doi: 10.1136/amiajnl-2012-001089] [Medline: 23011124]
6. Lin C, Payne TH, Nichol WP, Hoey PJ, Anderson CL, Gennari JH. Evaluating clinical decision support systems: monitoring CPOE order check override rates in the Department of Veterans Affairs' computerized patient record system. J Am Med Inform Assoc 2008 Sep 01;15(5):620-626. [doi: 10.1197/jamia.m2453]
7. Taylor LK, Tamblyn R. Reasons for physician non-adherence to electronic drug alerts. Stud Health Technol Inform 2004;107(Pt 2):1101-1105. [Medline: 15360983]
8. Hsieh TC, Kuperman GJ, Jaggi T, Hojnowski-Diaz P, Fiskio J, Williams DH, et al. Characteristics and consequences of drug allergy alert overrides in a computerized physician order entry system. J Am Med Inform Assoc 2004 Nov 01;11(6):482-491. [doi: 10.1197/jamia.m1556]
9. Shah N, Seger A, Seger D, Fiskio J, Kuperman G, Blumenfeld B. Improving acceptance of computerized prescribing alerts in ambulatory care. J Am Med Inform Assoc 2006;13(1):5-11. [doi: 10.1197/jamia.m1868]
10. Slight S, Seger D, Franz C, Wong A, Bates D. The national cost of adverse drug events resulting from inappropriate medication-related alert overrides in the United States. J Am Med Inform Assoc 2018 Sep 01;25(9):1183-1188 [FREE Full text] [doi: 10.1093/jamia/ocy066] [Medline: 29939271]
11. Nanji K, Seger D, Slight S, Amato M, Beeler P, Her Q, et al. Medication-related clinical decision support alert overrides in inpatients. J Am Med Inform Assoc 2018 May 01;25(5):476-481 [FREE Full text] [doi: 10.1093/jamia/ocx115] [Medline: 29092059]
12. Moltu C, Stefansen J, Svisdahl M, Veseth M. Negotiating the coresearcher mandate-service users' experiences of doing collaborative research on mental health. Disabil Rehabil 2012;34(19):1608-1616. [doi: 10.3109/09638288.2012.656792] [Medline: 22489612]
13. Paterno MD, Maviglia SM, Gorman PN, Seger DL, Yoshida E, Seger AC, et al. Tiering drug-drug interaction alerts by severity increases compliance rates. J Am Med Inform Assoc 2009 Jan 01;16(1):40-46. [doi: 10.1197/jamia.m2808]
14. Embi PJ, Leonard AC. Evaluating alert fatigue over time to EHR-based clinical trial alerts: findings from a randomized controlled study. J Am Med Inform Assoc 2012 Jun 01;19(e1):e145-e148 [FREE Full text] [doi: 10.1136/amiajnl-2011-000743] [Medline: 22534081]
15. Cash J. Alert fatigue. Am J Health Syst Pharm 2009 Dec 01;66(23):2098-2101. [doi: 10.2146/ajhp090181] [Medline: 19923309]

16. Baysari M, Tariq A, Day R, Westbrook J. Alert override as a habitual behavior: a new perspective on a persistent problem. J Am Med Inform Assoc 2017 Mar 01;24(2):409-412 [FREE Full text] [doi: 10.1093/jamia/ocw072] [Medline: 27274015]
17. Miller KJ, Shenhav A, Ludvig EA. Habits without values. Psychol Rev 2019 Mar;126(2):292-311 [FREE Full text] [doi: 10.1037/rev0000120] [Medline: 30676040]
18. Zhang C, Vanschoren J, van Wissen A, Lakens D, de Ruyter B, IJsselsteijn W. Theory-based habit modeling for enhancing behavior prediction. ArXiv Preprint posted online Jan 5, 2021. [FREE Full text]
19. Verplanken B, Orbell S. Reflections on past behavior: a self-report index of habit strength. J Appl Soc Psychol 2003;33(6):1313-1330. [doi: 10.1111/j.1559-1816.2003.tb01951.x]
20. Kahneman D. Thinking, Fast and Slow. New York City: Macmillan; 2011.
21. Baysari MT, Westbrook JI, Richardson KL, Day RO. The influence of computerized decision support on prescribing during ward-rounds: are the decision-makers targeted? J Am Med Inform Assoc 2011 Nov 01;18(6):754-759 [FREE Full text] [doi: 10.1136/amiajnl-2011-000135] [Medline: 21676939]
22. O'Reilly RC, Frank MJ. Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. Neural Comput 2006 Feb 01;18(2):283-328. [doi: 10.1162/089976606775093909] [Medline: 16378516]
23. Ashby FG, Turner BO, Horvitz JC. Cortical and basal ganglia contributions to habit learning and automaticity. Trends Cogn Sci 2010 May;14(5):208-215 [FREE Full text] [doi: 10.1016/j.tics.2010.02.001] [Medline: 20207189]
24. Ashby FG, Ennis JM, Spiering BJ. A neurobiological theory of automaticity in perceptual categorization. Psychol Rev 2007;114(3):632-656. [doi: 10.1037/0033-295x.114.3.632]
25. Tobias R. Changing behavior by memory aids: a social psychological model of prospective memory and habit development tested with dynamic field data. Psychol Rev 2009 Apr;116(2):408-438. [doi: 10.1037/a0015512] [Medline: 19348548]
26. Klein M, Mogles N, Treur J, Van Wissen A. A computational model of habit learning to enable ambient support for lifestyle change. 2011. URL: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.228.7928&rep=rep1&type=pdf [accessed 2022-01-10]
27. Baysari M, Westbrook J, Richardson K, Day R. Optimising computerised alerts within electronic medication management systems: a synthesis of four years of research. Stud Health Technol Inform 2014;204:1-6. [Medline: 25087519]
28. Zenziper Straichman Y, Kurnik D, Matok I, Halkin H, Markovits N, Ziv A, et al. Prescriber response to computerized drug alerts for electronic prescriptions among hospitalized patients. Int J Med Inform 2017 Nov;107:70-75. [doi: 10.1016/j.ijmedinf.2017.08.008] [Medline: 29029694]
29. Tamblyn R, Reidel K, Patel V. Physicians' response to computerised alerts for psychotropic drugs in older persons: a multilevel analysis of the associated alert, patient and physician characteristics. BMJ Open 2012 Sep 27;2(5):e001384 [FREE Full text] [doi: 10.1136/bmjopen-2012-001384] [Medline: 23024254]
30. Limayem, Hirt, Cheung. How habit limits the predictive power of intention: the case of information systems continuance. MIS Quarterly 2007;31(4):705. [doi: 10.2307/25148817]
31. Wood W, Tam L, Witt MG. Changing circumstances, disrupting habits. J Pers Soc Psychol 2005 Jun;88(6):918-933. [doi: 10.1037/0022-3514.88.6.918] [Medline: 15982113]
32. Polites GL, Karahanna E. The embeddedness of information systems habits in organizational and individual level routines: development and disruption. MIS Quarterly 2013 Jan 1;37(1):221-246. [doi: 10.25300/misq/2013/37.1.10]

Abbreviations

EMR: electronic medical record
IDC: indwelling catheter

Edited by T Leung; submitted 09.08.20; peer-reviewed by M Baysari, P Beeler, C Zhang; comments to author 18.11.20; revised version received 10.04.21; accepted 10.12.21; published 16.02.22.

Please cite as:
Wang L, Goh KH, Yeow A, Poh H, Li K, Yeow JJL, Tan G, Soh C
Habit and Automaticity in Medical Alert Override: Cohort Study
J Med Internet Res 2022;24(2):e23355
URL: https://www.jmir.org/2022/2/e23355
doi:10.2196/23355
PMID:35171102

Original Paper

SMART COVID Navigator, a Clinical Decision Support Tool for COVID-19 Treatment: Design and Development Study

Varun Suraj[1]; Catherine Del Vecchio Fitz[2], PhD; Laura B Kleiman[2], PhD; Suresh K Bhavnani[3], PhD, MArch; Chinmay Jani[4], MBBS; Surbhi Shah[5], MBBS; Rana R McKay[6], MD; Jeremy Warner[7*], MD; Gil Alterovitz[8*], PhD

[1]Biomedical Cybernetics Laboratory, Brigham and Women's Hospital, Boston, MA, United States

[2]Reboot Rx, Boston, MA, United States

[3]Preventive Medicine and Population Health Institute for Translational Sciences, University of Texas Medical Branch, University of Texas Health Science Center in Houston, Houston, TX, United States

[4]Department of Internal Medicine, Mount Auburn Hospital, Harvard Medical School, Cambridge, MA, United States

[5]Hematology, Oncology and Bone Marrow Transplantation, Mayo Clinic, Phoenix, AZ, United States

[6]Department of Medicine and Urology, University of California, San Diego, CA, United States

[7]Medicine and Biomedical Informatics, Vanderbilt University, Nashville, TN, United States

[8]Biomedical Cybernetics Laboratory, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States

[*]these authors contributed equally

Corresponding Author:
Jeremy Warner, MD
Medicine and Biomedical Informatics
Vanderbilt University
2525 West End Ave, Suite 1500
Nashville, TN, 37203
United States
Phone: 1 615 936 3524
Email: jeremy.warner@vumc.org

Abstract

Background: COVID-19 caused by SARS-CoV-2 has infected 219 million individuals at the time of writing of this paper. A large volume of research findings from observational studies about disease interactions with COVID-19 is being produced almost daily, making it difficult for physicians to keep track of the latest information on COVID-19's effect on patients with certain pre-existing conditions.

Objective: In this paper, we describe the creation of a clinical decision support tool, the SMART COVID Navigator, a web application to assist clinicians in treating patients with COVID-19. Our application allows clinicians to access a patient's electronic health records and identify disease interactions from a large set of observational research studies that affect the severity and fatality due to COVID-19.

Methods: The SMART COVID Navigator takes a 2-pronged approach to clinical decision support. The first part is a connection to electronic health record servers, allowing the application to access a patient's medical conditions. The second is accessing data sets with information from various observational studies to determine the latest research findings about COVID-19 outcomes for patients with certain medical conditions. By connecting these 2 data sources, users can see how a patient's medical history will affect their COVID-19 outcomes.

Results: The SMART COVID Navigator aggregates patient health information from multiple Fast Healthcare Interoperability Resources–enabled electronic health record systems. This allows physicians to see a comprehensive view of patient health records. The application accesses 2 data sets of over 1100 research studies to provide information on the fatality and severity of COVID-19 for several pre-existing conditions. We also analyzed the results of the collected studies to determine which medical conditions result in an increased chance of severity and fatality of COVID-19 progression. We found that certain conditions result in a higher likelihood of severity and fatality probabilities. We also analyze various cancer tissues and find that the probabilities for fatality vary greatly depending on the tissue being examined.

Conclusions: The SMART COVID Navigator allows physicians to predict the fatality and severity of COVID-19 progression given a particular patient's medical conditions. This can allow physicians to determine how aggressively to treat patients infected with COVID-19 and to prioritize different patients for treatment considering their prior medical conditions.

(J Med Internet Res 2022;24(2):e29279) doi:10.2196/29279

KEYWORDS

COVID-19; clinical decision support; precision medicine; web application; FHIR

Introduction

Precision medicine, as defined by the National Institutes of Health's Precision Medicine Initiative, is "an emerging approach to disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person" [1]. Recent increases in the availability of electronic patient data have facilitated the development of precision medicine. For instance, electronic health records (EHRs) store all information collected in hospitals, such as blood tests, radiographs, diagnostic tests, and any biographical information about a patient, such as their age, weight, or height [2]. This approach to medical treatment could prove to be useful in dealing with the COVID-19 caused by SARS-CoV-2, which has been ravaging the United States and the world [3,4]. Currently, a vast amount of research is being conducted to understand how a patient's underlying health conditions interact with the progression of the virus infection. For example, certain conditions such as diabetes and heart disease have been found to raise the severity and fatality rates of patients who have become infected with the virus [5]. Applying the tools of precision medicine can help doctors customize treatment for patients affected by COVID-19 based on the patient's underlying conditions. Attempts at clinical decision support systems for COVID-19 have been made using information on risk factors and biomarker measurements [6,7].

With the rapid spread of COVID-19 and the scarcity of physician resources and time, there is an immediate need for a clinical decision support system that provides patient and disease interaction information to clinicians to allow them to practice precision medicine. In this paper, we describe the creation of the SMART COVID Navigator, a web-based application designed to assist clinicians to relate patient risk factors to the growing amount of research that identifies how various underlying patient conditions affect the progression of COVID-19. This application identifies patient risk factors based on integrating comprehensive data available through multiple EHRs. The application then allows clinicians to quickly access a large set of research studies based on their patient's medical history in an easy-to-access format, helping promote updated information on COVID-19 and its risks for a diverse community. This tool will simplify a clinician's search for relevant research and findings and support clinical treatment.

The SMART COVID Navigator connects patient information from multiple EHR servers to 2 databases of COVID-19 research studies. This will allow clinicians to access data-driven research based on a particular patient's risk factors. The SMART COVID Navigator builds on the framework developed by the SMART Cancer Navigator, which offers clinical decision support by connecting patient EHR information to cancerous gene variants [8,9]. The Navigator is a further step for creating Fast Healthcare Interoperability Resources (FHIR) protocol–based tools to support personalized medicine [10]. This web-based application was built by researchers at the Biomedical Cybernetics Laboratory in Harvard Medical School. It was created using an Angular and Bootstrap front-end framework. The code is available at [11].

Methods

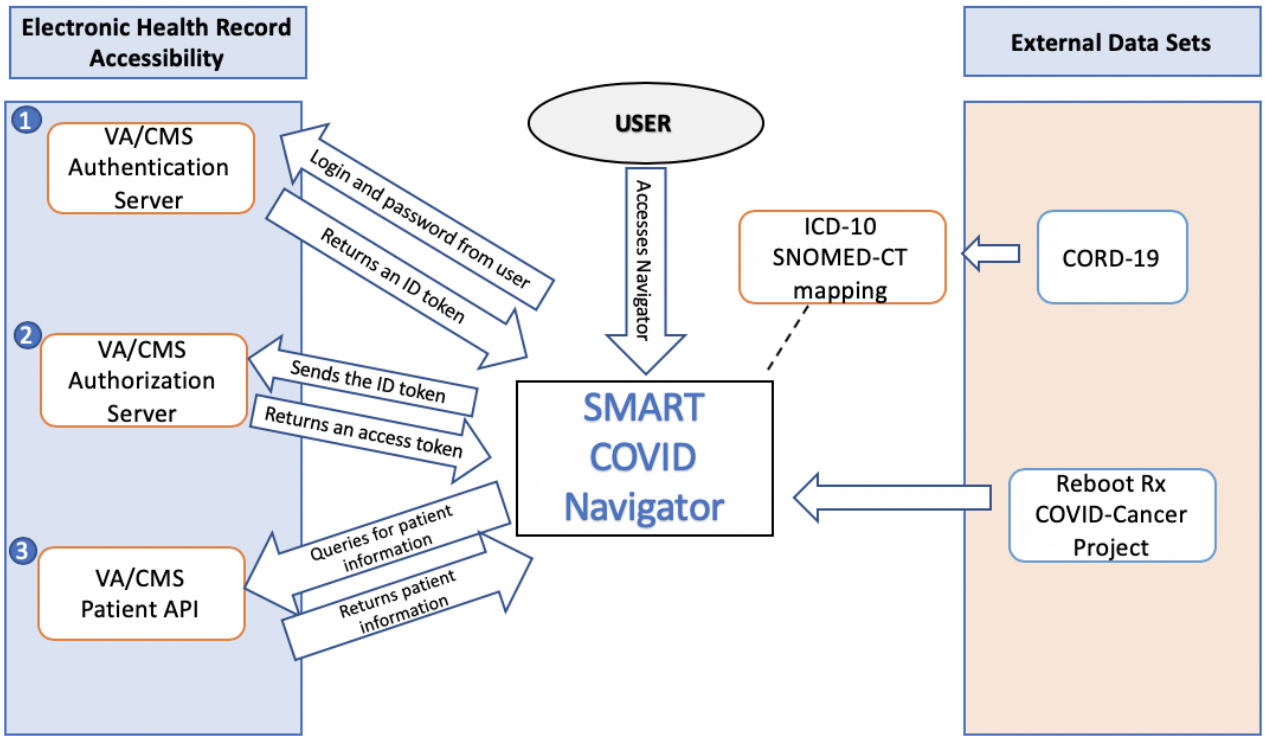
Accessing Multiple FHIR Servers

The SMART COVID Navigator allows the user to log into 2 EHRs: the Veterans Affairs (VA) and the Center for Medicare and Medicaid Services (CMS). There are significant advantages to the application being able to log into multiple EHR servers [12]. Doctors viewing their patient's medical information through the application have the ability to discern any discrepancies between the data sources. Similarly, doctors and patients can be sure that they are receiving the most up-to-date information regarding their health records, as any information not captured in one of the EHRs would most likely be shown in the other. The VA and CMS servers follow the FHIR standard [13]. The FHIR platform allows for the interoperability of the navigator with other EHRs that follow the FHIR protocol. The application is registered with the relevant EHRs. The system can be expanded to access additional health records, provided their application programming interfaces (APIs) follow the FHIR standard. The EHRs provide the application with a client ID and a client secret, which is used for authentication purposes. The following section describes the login process.

To log into the 2 API-enabled EHR systems, the application implements the OAuth2 [14] and OpenID Connect [15] standards to achieve secure authentication and authorization. The authentication process involves a user entering their login credentials, while the authorization process requires a request to the EHR's server to obtain an access token. Upon receiving the access token, the application is authorized to retrieve relevant patient information. The left side of Figure 1 depicts the system architecture for the EHR access process. The login process begins when the user clicks one of the login buttons. The user is redirected to the login portal for the EHR of their choice to enter their login credentials. The OAuth2 process redirects the user back to the application with an ID token in the URL (step 1 in Figure 1). With this ID token, the application then requests an access token from the relevant API to obtain access to the patient's medical information. The application accomplishes this second step by sending the ID token along with additional information such as the application's client ID and client secret back to the EHR server through an HTML POST request. The

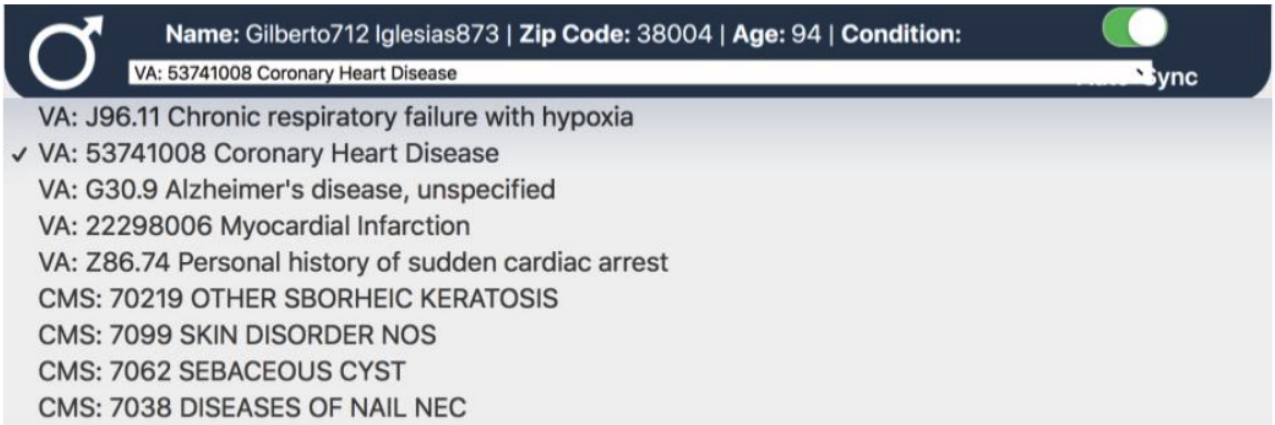
server then returns an access token, which the application can use to gain access to any part of the patient's profile (step 2 in Figure 1). The access token is saved to the local storage of the app; thus, if the user refreshes the application or attempts to log into another EHR, they will still retain access to the first EHR. This 2-step process ensures better security of sensitive medical data. From the EHRs, the following information is retrieved: (1) the patient's name, (2) their location zip code, (3) their date of birth (which is used to calculate their age), and (4) a list of medical conditions associated with the patient along with a numerical code presenting the condition (step 3 in Figure 1).

Figure 1. Architecture diagram of the SMART COVID Navigator. API: application programming interface; CMS: Center for Medicare and Medicaid Services; CORD-19: COVID-19 Open Research Dataset; ICD-10: International Classification of Diseases-tenth version; SNOMED-CT: SNOMED Clinical Terms; VA: Veterans Affairs.



The EHR server results are visually displayed at the top of the screen. This menu appears when a user logs into either one of the servers. It displays demographic information such as the patient's name, age, and current zip code. It also shows a list of medical conditions retrieved from the patient's profile with the list of conditions appearing in a dropdown menu (Figure 2). If the user is logged into both the VA and CMS servers, then the application retrieves demographic information from the VA server, while the condition list is a combined list from both EHRs. Please note that the patient information displayed in Figure 2 and in other figures below are based on information taken from VA- and CMS-simulated patient data; since these data are not from a real patient, there are no Health Insurance Portability and Accountability Act concerns. With these patient conditions now available in the Navigator, the next step is to connect these conditions to risk factors shown to affect COVID-19 progression.

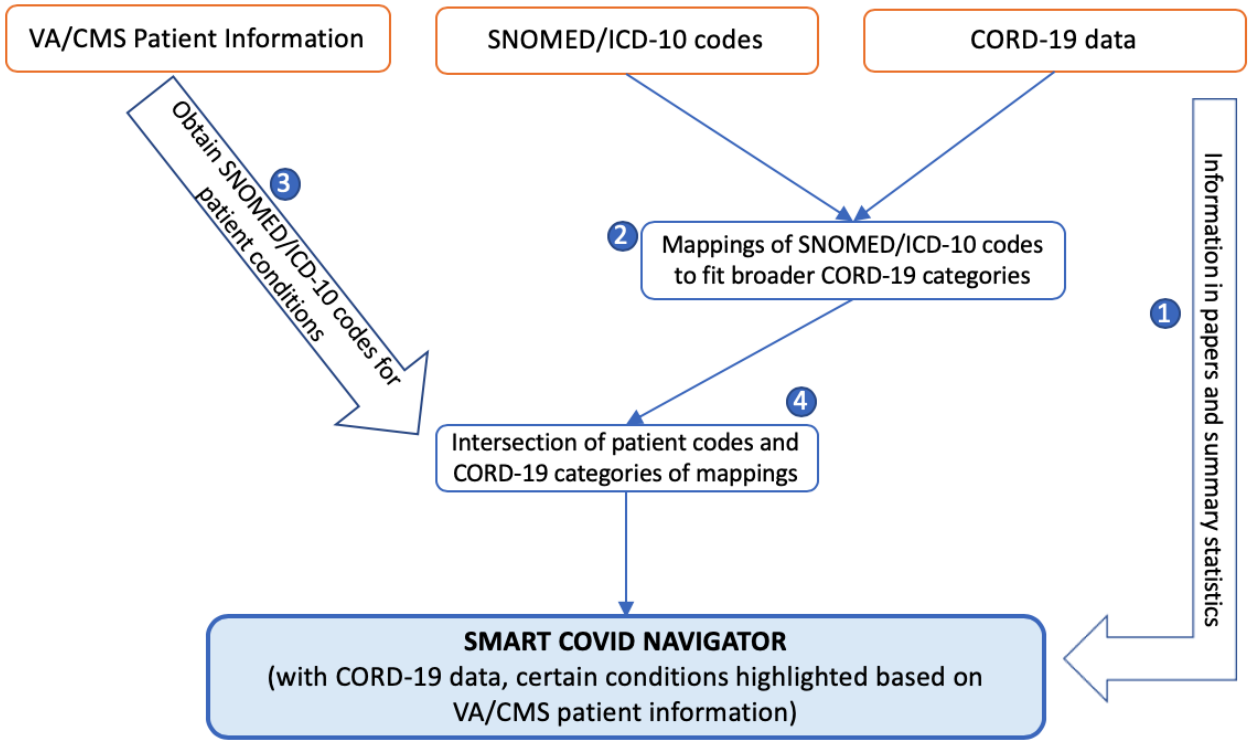
Figure 2. The disease condition list for a Veterans Affairs test patient. This dropdown appears from the patient header.



Accessing Artificial Intelligence–Powered Knowledge Base From the COVID-19 Open Research Dataset

The SMART COVID Navigator collects risk factor information from 2 external data sets, as shown on the right side of Figure 1 and relates it to patient medical conditions. The first data set is the COVID-19 Open Research Dataset (CORD-19), created by the White House and a set of leading research groups such as the Georgetown University's Center for Security and Emerging Technology and the National Library of Medicine, National Institutes of Health [16]. This resource currently contains over 200,000 scholarly papers related to COVID-19. Researchers apply artificial intelligence techniques to create this knowledge base. CORD-19 sources papers from the World Health Organization, PubMed Central, bioRxiv, and medRxiv. These can be in the form of physical printouts, PDF files, or XML files. The papers and preprints are then collected by Semantic Scholar [17], and the resulting metadata are harmonized and deduplicated. The full text of the papers is then extracted [18]. This knowledge base is stored on Kaggle, a web-based community of data scientists and machine learning practitioners [19]. The Navigator utilizes the "risk factors" section of CORD-19 to support clinical assessment of how a patient's medical conditions affect their chances of having a severe or fatal COVID-19 infection (step 1 in Figure 3).

Figure 3. Architecture diagram of the interactions between the Veterans Affairs and Center for Medicare and Medicaid Services electronic health record systems, SNOMED/International Classification of Diseases-tenth version codes, and the COVID-19 Open Research Dataset. CMS: Center for Medicare and Medicaid Services; CORD-19: COVID-19 Open Research Dataset; ICD-10: International Classification of Diseases-tenth version; SNOMED-CT: SNOMED Clinical Terms; VA: Veterans Affairs.



The CORD-19 data are publicly available and are stored in CSV files. Each file tracks studies relating to a specific condition (eg, hypertension, heart disease) that is being tracked. Each row in the CSV file represents a study. The study name, link, date, and significance of the severity and fatality of the patients are all provided. Textbox 1 provides a list of the risk factors provided in CORD-19. As shown in Textbox 1, the CORD-19 data set consists of 28 risk factors—22 of which are disease conditions, while the other 6 are patient biographical characteristics. The comprehensive clinical terminology coding systems, SNOMED Clinical Terms used by the VA [20] and the International Classification of Diseases-tenth version (ICD-10) used by CMS [21], consist of thousands of more specific disease identifiers than the broader disease categories tracked by CORD-19. We mapped the identifiers in the SNOMED Clinical Terms and ICD-10 code sets to the CORD-19 risk factors to be able to match patient conditions to disease categories tracked by CORD-19 (step 2 in Figure 3). This mapping was accomplished through word matching. For each of the 28 risk factors present in CORD-19, the name of the risk factor or a part of the name was matched with every instance of that name occurring in the SNOMED Clinical Terms and ICD-10 code sets by using a Python script created by the authors. For instance, the risk factor "diabetes" links to any disease classification in the code sets that has the word "diabetes" or variations such as "diabetic." Because the disease identifiers in the code sets are more specific than the broad risk factors tracked by CORD-19, many risk factors align with hundreds of more specific disease names from the code sets. Since the mapping involved matching words or parts of words using Python code (and was not done manually), we believe there is unlikely to be incorrect mapping. For example, we searched for the string "diabet," which would capture both "diabetes" and "diabetic." Nevertheless, it is possible that some conditions in the 2 code sets might not have been captured despite our attempts at careful mapping.

Textbox 1. Risk factors tracked by the COVID-19 Open Research Dataset.

Risk factors

- Age
- Endocrine diseases
- Asthma
- Ethnicity: Hispanic versus non-Hispanic
- Autoimmune disorders
- Heart disease
- Chronic obstructive pulmonary disease
- Heart failure
- Cancer
- Hypertension
- Cardiovascular and cerebrovascular disease
- Immune system disorders
- Cerebrovascular disease
- Male gender
- Chronic digestive disorders
- Neurological disorders
- Chronic kidney disease
- Overweight or obese
- Chronic liver disease
- Race: Asian versus White
- Chronic respiratory diseases
- Race: Black versus White
- Dementia
- Race: Other versus White
- Diabetes
- Respiratory system diseases
- Drinking
- Smoking status

Visually in the Navigator, the information from CORD-19 regarding studies related to COVID-19 risk factors appears as a menu of conditions (Figure 4). When the user clicks a particular condition, a pop-up screen is displayed, providing all the information collected about studies relevant to that condition (Figure 5). This will allow a user to view the studies and use the findings for clinical decision support. The study name (with a hyperlink to the full study), date, and the significance of the severity and fatality statistics are shown for each study. At the top of the pop-up, an overview of the studies associated with the chosen condition is displayed. The information included are the total number of studies for that condition, the percentage of studies that found the selected condition to cause a significant change in the severity of COVID-19 progression (out of the studies that measure for severity), and the percentage of studies that found a significant change in fatality due to COVID-19. We alert the user if the proportion of papers finding a significant result is greater than 50%. If a particular risk factor is found to be associated with the severity or fatality due to COVID-19 infection in a high proportion of observational studies, then physicians should pay more attention to patients with that risk factor. However, individual physicians should make their own decisions based on the information in the SMART COVID Navigator and the severity of their patients' health conditions.

Figure 4. A view of the application logged into a test Veterans Affairs profile. A button appears for each condition that is being tracked. Based on the conditions in the patient's profile, certain risk factors on the screen are highlighted in orange.

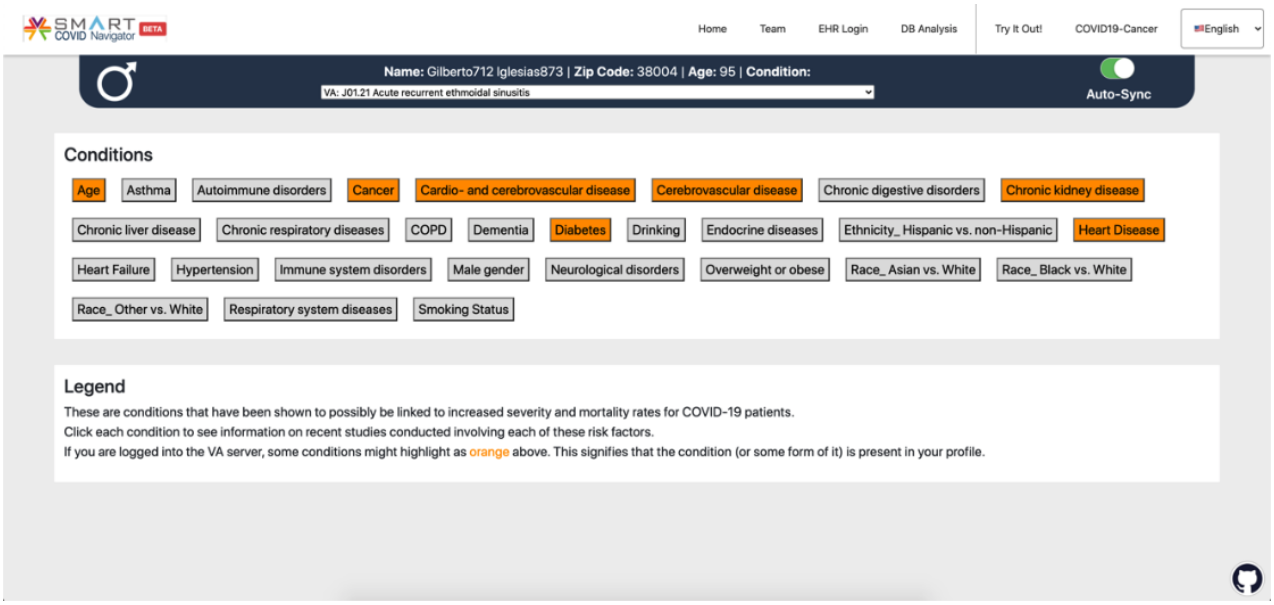
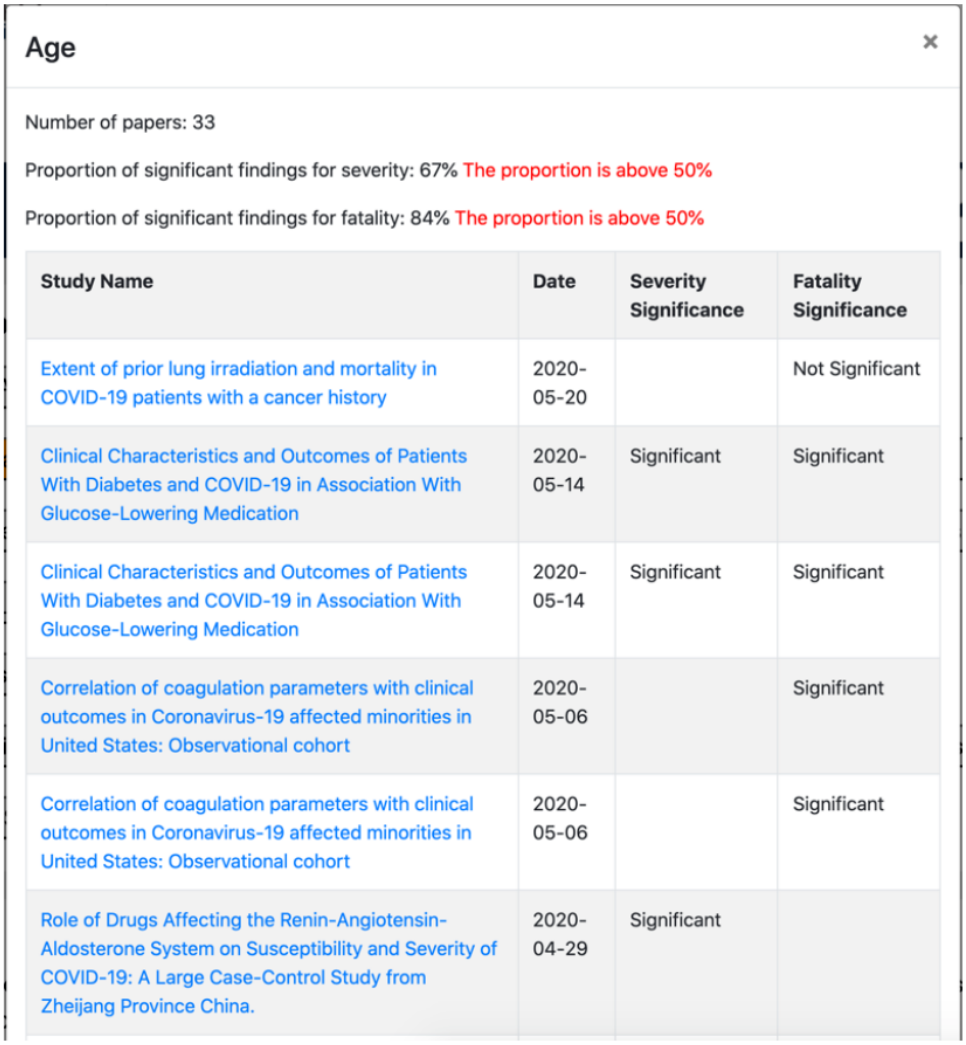


Figure 5. Presentation of information from the COVID-19 Open Research Dataset relating to a particular risk factor (in this case, the example is for the "age" risk factor).



Age

Number of papers: 33

Proportion of significant findings for severity: 67% The proportion is above 50%

Proportion of significant findings for fatality: 84% The proportion is above 50%

Study Name	Date	Severity Significance	Fatality Significance
Extent of prior lung irradiation and mortality in COVID-19 patients with a cancer history	2020-05-20		Not Significant
Clinical Characteristics and Outcomes of Patients With Diabetes and COVID-19 in Association With Glucose-Lowering Medication	2020-05-14	Significant	Significant
Clinical Characteristics and Outcomes of Patients With Diabetes and COVID-19 in Association With Glucose-Lowering Medication	2020-05-14	Significant	Significant
Correlation of coagulation parameters with clinical outcomes in Coronavirus-19 affected minorities in United States: Observational cohort	2020-05-06		Significant
Correlation of coagulation parameters with clinical outcomes in Coronavirus-19 affected minorities in United States: Observational cohort	2020-05-06		Significant
Role of Drugs Affecting the Renin-Angiotensin-Aldosterone System on Susceptibility and Severity of COVID-19: A Large Case-Control Study from Zhejiang Province China.	2020-04-29	Significant	

The patient's biographical and health information retrieved from the EHRs can be used to better guide clinicians on the health conditions they should be concerned about in relation to COVID-19. The Navigator uses the previously described mappings from SNOMED Clinical Terms and ICD-10 identifiers to the CORD-19 risk factors; wherever there is a match between the code in a patient's EHR condition list (step 3 of Figure 3) and those associated with each CORD-19 risk factor, the condition highlights orange (step 4 of Figure 3), thereby alerting the user to presence of the condition of concern (Figure 4). If the patient is older than 60 years, then the "age" condition is highlighted as well. The clinician should click on any highlighted conditions to access further information from the CORD-19 database, as this information is likely to be relevant for the patient in question.

COVID-19 in Patients With Cancer

Cancer has been shown to be a risk factor for COVID-19, and the different types of cancers affect COVID-19 progression differently [22]; therefore, it is important for physicians to understand disease progression if patients with cancer become infected with COVID-19. To help with this, the SMART COVID Navigator accesses Reboot Rx's Reboot: COVID-Cancer Project data sets in order to display information regarding the impact of COVID-19 and its treatment on patients with cancer [23]. The Reboot: COVID-Cancer Project identifies relevant published clinical studies and extracts and aggregates the data from those studies into 2 data sets. The first data set examines COVID-19 disease progression for patients with cancer. The second data set examines what effect drugs currently being tested for COVID-19 treatment could have on cancer (independent of their effect on COVID-19). The data sets are publicly available on the Reboot: COVID-Cancer Project website via interaction dashboards and can be requested in the form of Excel files. This feature gives an organized tool for clinicians to better understand COVID-19 outcomes for different types of cancer. Textbox 2 shows each of the tissue types that are examined in the Reboot: COVID-Cancer Project data sets (the numbers in parentheses indicate which data sets the tissue type is in). As shown in Textbox 2, the data sets consist of 30 tissue types. Tissue types were chosen to distinguish the cancers instead of cancer types, as the number of tissue types was more manageable and therefore would be easier to access for the user of the application.

When a user navigates to the "COVID19-Cancer" tab of the SMART COVID Navigator, they will see a menu of tissue types, similar to how the CORD-19 risk factors are displayed, as well as another button that displays the summary information for all of the tissues in 1 pop-up, colored blue to distinguish it from the other buttons representing tissue types (Figure 6). When one of the buttons is clicked, a pop-up appears with 2 tabs. The first tab displays information retrieved from data set 1 regarding patient outcomes, and the second tab displays information from data set 2 about COVID-19 drugs that could be useful for cancer treatment (Figure 7). When the summary tab is clicked, information retrieved from data set 1 about all tissue types in that data set is displayed, so that clinicians can easily see all the patient outcome statistics on 1 central page (Figure 8). Please note that the use of data set 1 in the SMART COVID Navigator is to display patient outcomes for the treatment of COVID-19 in patients with cancer. This use is similar to how we used the CORD-19 data set. The display of data set 2 is for a different purpose—to inform physicians treating cancer that how drugs that are being used for COVID-19 treatment have an application in cancer treatment.

Textbox 2. Tissue types tracked by the Reboot: COVID-Cancer Project data sets. The numbers in parentheses indicate which data sets the tissue type is in.

Tissue types

- Bladder/unitary tract (1,2)
- Head and neck (1,2)
- Pancreas (2)
- Bone and soft tissue (1)
- Hematologic not specified (1)
- Pleura (1)
- Bowel (1,2)
- Kidney (1,2)
- Prostate (1,2)
- Bowel, esophagus/stomach (2)
- Liver (1,2)
- Sarcoma (2)
- Brain/central nervous system (1,2)
- Lung (1,2)
- Skin (1,2)
- Breast (1,2)
- Lymphoid (1,2)
- Soft tissue (2)
- Cervix (1)
- Lymphoid, myeloid (2)
- Testis (2)
- Esophagus/stomach (2)
- Myeloid (1,2)
- Thoracic (1)
- Genitourinary (1)
- Not specified (1,2)
- Thymus (1,2)
- Gynecological (1)
- Ovary/fallopian tube (1,2)
- Thyroid (2)

Figure 6. Standard view of the COVID-19 cancer tab of the COVID Navigator. The "Summary" button is highlighted in blue to distinguish it from the other buttons that represent tissue types.

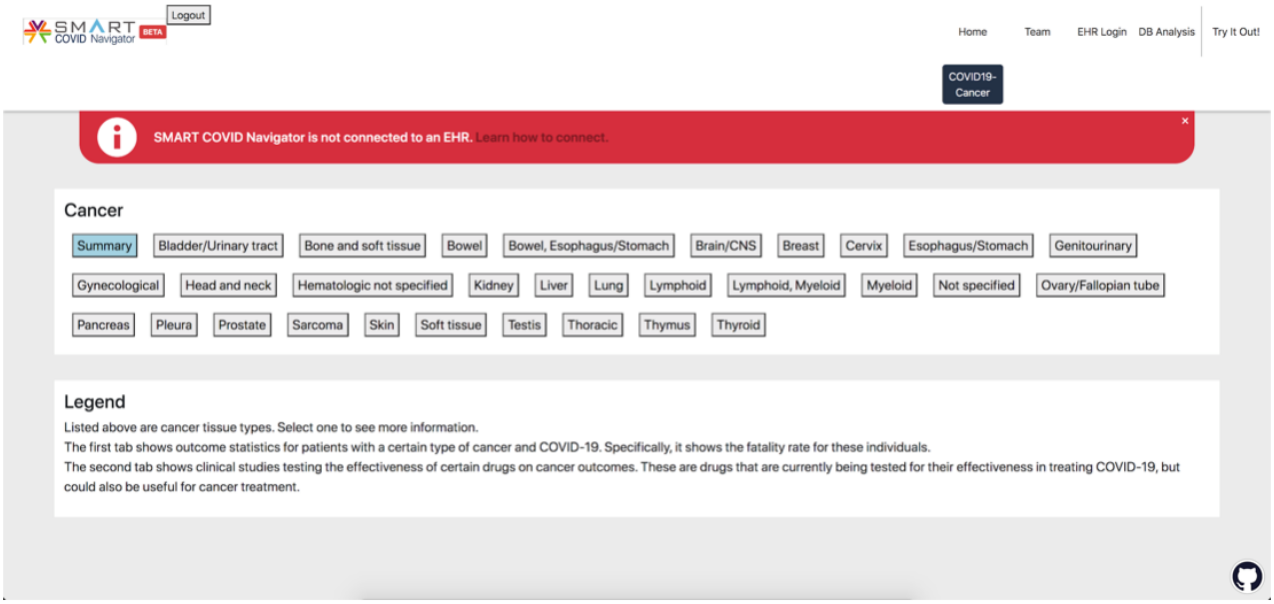
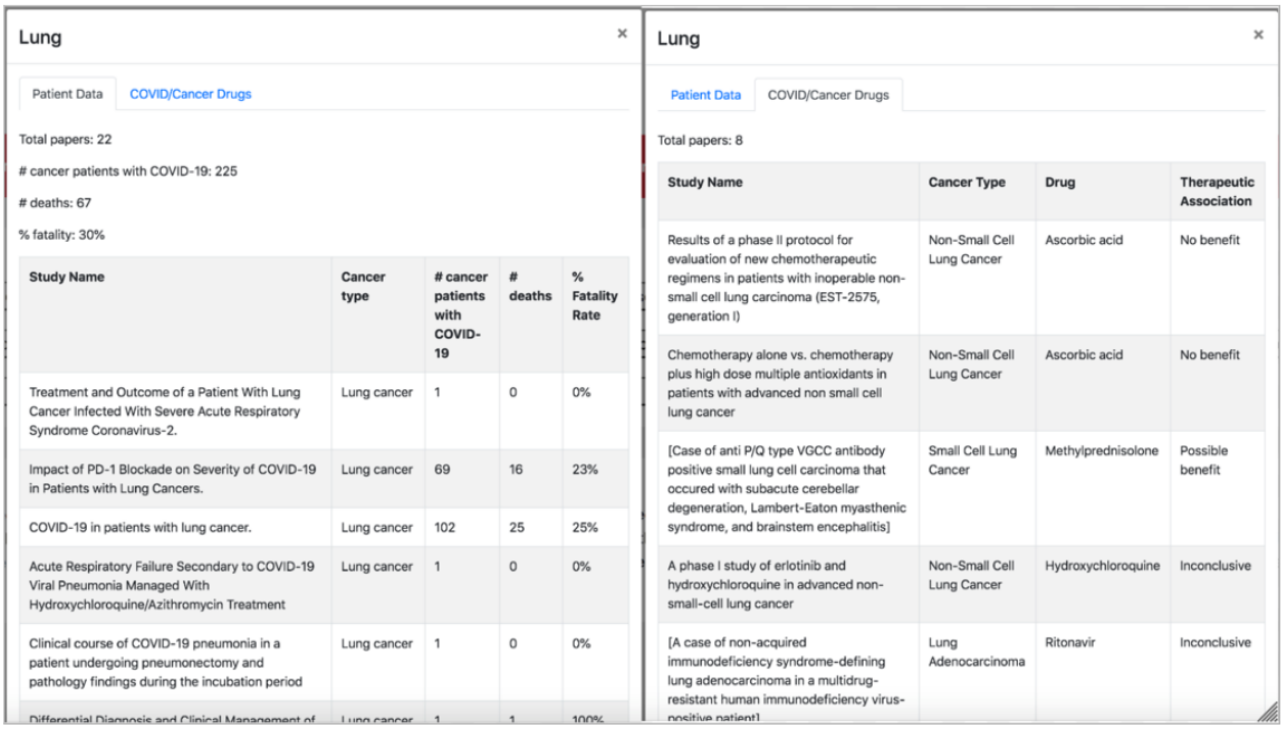


Figure 7. When a tissue is clicked, a pop-up appears with 2 tabs (this figure shows the pop-up for lung tissue). The first tab (on the left) tracks patient outcomes for patients with lung cancer who contract COVID-19. The second tab (on the right) tracks drugs being tested for COVID-19 that could have an impact on cancer of the selected tissue.



Lung

Patient Data | COVID/Cancer Drugs

Total papers: 22

cancer patients with COVID-19: 225

deaths: 67

% fatality: 30%

Study Name	Cancer type	# cancer patients with COVID-19	# deaths	% Fatality Rate
Treatment and Outcome of a Patient With Lung Cancer Infected With Severe Acute Respiratory Syndrome Coronavirus-2.	Lung cancer	1	0	0%
Impact of PD-1 Blockade on Severity of COVID-19 in Patients with Lung Cancers.	Lung cancer	69	16	23%
COVID-19 in patients with lung cancer.	Lung cancer	102	25	25%
Acute Respiratory Failure Secondary to COVID-19 Viral Pneumonia Managed With Hydroxychloroquine/Azithromycin Treatment	Lung cancer	1	0	0%
Clinical course of COVID-19 pneumonia in a patient undergoing pneumonectomy and pathology findings during the incubation period	Lung cancer	1	0	0%
Differential Diagnosis and Clinical Management of	Lung cancer	1	1	100%

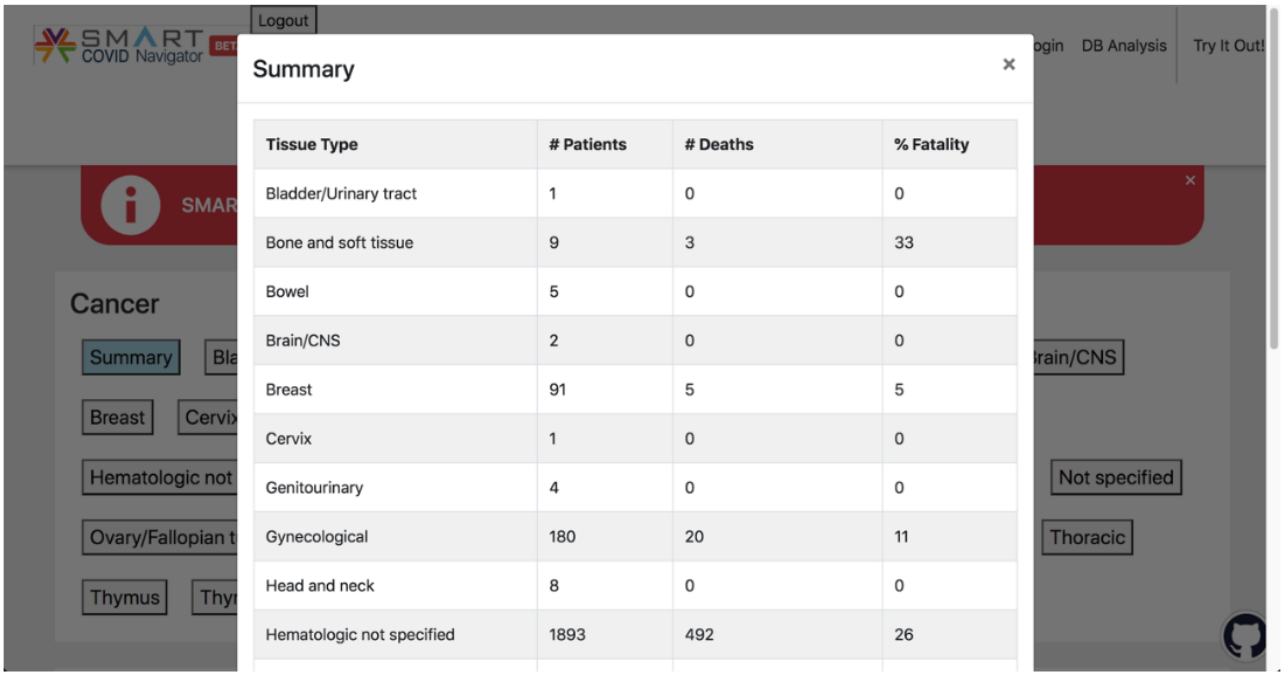
Lung

Patient Data | COVID/Cancer Drugs

Total papers: 8

Study Name	Cancer Type	Drug	Therapeutic Association
Results of a phase II protocol for evaluation of new chemotherapeutic regimens in patients with inoperable non-small cell lung carcinoma (EST-2575, generation I)	Non-Small Cell Lung Cancer	Ascorbic acid	No benefit
Chemotherapy alone vs. chemotherapy plus high dose multiple antioxidants in patients with advanced non small cell lung cancer	Non-Small Cell Lung Cancer	Ascorbic acid	No benefit
[Case of anti P/Q type VGCC antibody positive small lung cell carcinoma that occured with subacute cerebellar degeneration, Lambert-Eaton myasthenic syndrome, and brainstem encephalitis]	Small Cell Lung Cancer	Methylprednisolone	Possible benefit
A phase I study of erlotinib and hydroxychloroquine in advanced non-small-cell lung cancer	Non-Small Cell Lung Cancer	Hydroxychloroquine	Inconclusive
[A case of non-acquired immunodeficiency syndrome-defining lung adenocarcinoma in a multidrug-resistant human immunodeficiency virus-positive patient]	Lung Adenocarcinoma	Ritonavir	Inconclusive

Figure 8. When the blue "Summary" button is clicked, this pop-up appears, which displays the patient outcome summary statistics for each tissue type in a concise manner.



Tissue Type	# Patients	# Deaths	% Fatality
Bladder/Urinary tract	1	0	0
Bone and soft tissue	9	3	33
Bowel	5	0	0
Brain/CNS	2	0	0
Breast	91	5	5
Cervix	1	0	0
Genitourinary	4	0	0
Gynecological	180	20	11
Head and neck	8	0	0
Hematologic not specified	1893	492	26

Data Governance

There are 2 sources of data used in the application. The first is patient data from the VA and CMS, which are accessed when a patient logs into their accounts in these EHRs. The SMART COVID Navigator does not store any patient data. These data are accessed by the patient or physician under authorization by the patient. These data are used to identify patient risk factors when the physicians/patients are using the application. When they are done using the application and log out, the patient data are deleted from the local machine, which is being used to access the app. At no point is patient data stored in the main server running the application. The second source of data is from the publicly available CORD-19 and Reboot: COVID-Cancer Project data sets for use in the application. Given these are publicly available repositories of research papers, we do not implement any specific data governance with respect to this information.

Results

Knowledge Base Analysis

In addition to the development of the SMART COVID Navigator, we performed analyses on the research studies in the CORD-19 and the Reboot: COVID-Cancer Project knowledge bases. Given the large number of studies both in the CORD-19 knowledge base and the COVID-Cancer data sets, it would be difficult for a clinician to quickly assess if a particular patient's disease condition or cancer type interacts significantly with COVID-19. For the CORD-19 knowledge base, we explored whether on average, the studies found significant results for each of the 28 conditions examined, and for the Reboot: COVID-Cancer Project data sets, we explored the likelihood of mortality for a particular cancer type based on the aggregate data collected from the data sets.

CORD-19 Knowledge Base Analysis

First, we discuss the results from the CORD-19 knowledge base relating to risk factors. We found that out of the 816 studies examined in the risk factors section, 355 studies (43.5%) studied only how a particular disease condition affected the severity of COVID-19 progression, 328 studies (40.2%) examined only how the disease condition affected fatality, 89 studies (10.9%) examined both severity and fatality, and 44 studies (5.4%) had no entry for severity nor fatality. We examined each of the 28 individual risk conditions to identify which of them have been found more frequently to be a significant factor in COVID-19 severity and fatality rates. These data are displayed in Tables 1 and 2. Table 1 presents the data for each condition regarding severity, and Table 2 presents the data regarding fatality.

Table 1. Severity statistics for all risk factors.

Risk factor	Total papers (N)	Studies on COVID-19 severity (n)	Studies showing significant severity, n (%)[a]
Age	33	15	10 (67)
Asthma	6	3	0 (0)
Autoimmune disorders	3	1	0 (0)
Cancer	26	16	8 (50)
Cardiovascular and cerebrovascular disease	10	7	4 (57)
Cerebrovascular disease	20	8	6 (75)
Chronic digestive disorders	5	2	0 (0)
Chronic kidney disease	38	19	10 (53)
Chronic liver disease	12	6	1 (17)
Chronic respiratory diseases	29	12	6 (50)
Chronic obstructive pulmonary disease	41	23	18 (78)
Dementia	5	2	1 (50)
Diabetes	100	57	33 (58)
Drinking	1	1	0 (0)
Endocrine diseases	4	4	2 (50)
Ethnicity: Hispanic versus non-Hispanic	15	6	0 (0)
Heart disease	98	54	42 (78)
Heart failure	17	3	1 (33)
Hypertension	100	60	35 (58)
Immune system disorders	6	4	3 (75)
Male gender	100	59	29 (49)
Neurological disorders	6	3	1 (33)
Overweight or obese	43	28	18 (64)
Race: Asian versus White	5	2	0 (0)
Race: Black versus White	21	7	2 (29)
Race: Other versus White	9	3	0 (0)
Respiratory system diseases	7	6	5 (83)
Smoking status	56	33	16 (48)

[a]The proportion of studies showing significant severity was calculated from the number of studies showing severity of COVID-19 in people with the particular risk factor.

Table 2. Fatality statistics for all risk factors.

Risk factor	Total papers (N)	Studies on fatality (n)	Studies with significant fatality value, n (%)[a]
Age	33	19	16 (84)
Asthma	6	4	2 (50)
Autoimmune disorders	3	3	2 (67)
Cancer	26	13	6 (46)
Cardiovascular and cerebrovascular disease	10	3	2 (67)
Cerebrovascular disease	20	14	10 (71)
Chronic digestive disorders	5	2	0 (0)
Chronic kidney disease	38	23	14 (61)
Chronic liver disease	12	5	3 (60)
Chronic respiratory diseases	29	19	13 (68)
Chronic obstructive pulmonary disease	41	19	8 (42)
Dementia	5	4	4 (100)
Diabetes	100	46	26 (57)
Drinking	1	0	N/A[b]
Endocrine diseases	4	0	N/A
Ethnicity: Hispanic versus non-Hispanic	15	11	6 (55)
Heart disease	98	51	43 (84)
Heart failure	17	14	9 (64)
Hypertension	100	43	21 (49)
Immune system disorders	6	4	4 (100)
Male gender	100	39	20 (51)
Neurological disorders	6	3	3 (100)
Overweight or obese	43	20	12 (60)
Race: Asian versus White	5	4	3 (75)
Race: Black versus White	21	17	7 (41)
Race: Other versus White	9	9	3 (33)
Respiratory system diseases	7	2	2 (100)
Smoking status	56	26	5 (19)

[a]The proportion of studies showing significant fatality was calculated from the number of studies showing fatality of COVID-19 in people with the particular risk factor.

[b]N/A: not applicable.

Table 1 documents severity statistics; specifically, the number and proportion of studies that find if a given condition results in a statistically significant change in the likelihood of a patient having a severe COVID-19 infection. For example, out of the 98 studies on heart disease, 54 of them study severity and 42 (78%) found a significant result for severity. However, chronic liver disease seems to not have a significant impact on COVID-19 severity, as out of the 12 papers that study this condition, 6 study severity and only 1 study (17%) found a significant correlation. This application alerts users if more than 50% of papers report a significant finding.

Table 2 is similar to Table 1, except that instead of documenting findings about COVID-19 severity, it documents statistics regarding fatality. We can again look at some examples of medical conditions. As in the case of severity of COVID-19 progression, heart disease is a useful predictor of COVID-19 fatality; out of the 51 papers that measure fatality statistics for the condition, 43 of them (84%) found a significant correlation. Being a smoker, however, does not seem to have the same high degree of correlation; only 5 (19%) out of the 26 papers that measure fatality statistics for smoking status found a significant result. Similarly to that for severity, the application alerts users if more than 50% of papers report a significant finding for fatality.

There are conditions for which the papers tracked by CORD-19 do not offer a definitive answer. Cancer is one of these conditions; 50% (8/16) of studies found a significant finding for severity, while 46% (6/13) found a significant finding for

fatality. A possible explanation for this inconclusive result is the existence of various types of cancers in different tissues of the body, resulting in a mixed result when grouping various cancer types together. For such cases, clinicians will need to access additional information from other papers. As discussed before, we display the severity and fatality percentage summary statistics in the application for each condition so that clinicians can obtain a quick overview of the significance of that condition without having to review each study in detail. However, we caution that clinicians might want to review the studies in greater detail depending on the medical condition of the concerned patient.

Reboot: COVID-Cancer Project Knowledge Base Analysis

In addition to the CORD-19 analysis, we conducted an analysis based on the papers in the Reboot: COVID-Cancer Project data set relating to patient outcomes. This data set stores the number of patients with a specific type of cancer and with COVID-19, the number of those patients who died, and the percentage fatality found. The Navigator aggregates these data to offer the user summary statistics for each of the tissue types. Table 3 documents the aggregate patient outcomes for each cancerous tissue, compiled by adding the number of patients with COVID-19 and number of deaths reported in all the papers relating to the given tissue type. Some cancer tissue types in the table can be seen to be associated with a relatively high fatality rate. For instance, out of the 200 patients with thoracic cancer, 66 died (33%) due to COVID-19. Similarly, 152 out of the 504 patients with lymphoid cancer died (30.1%) and 67 of the 225 patients with lung cancer died (29.7%), indicating that cancers of these tissue types generally result in a higher fatality rate for COVID-19. However, only 5 out of the 91 patients with breast cancer died (6%), indicating that breast cancer is associated to a lesser degree with COVID-19 fatality than some other cancer tissue types.

For some tissue types, the data available through the Reboot: COVID-Cancer Project do not contain an adequate number of patients to support clinical decisions. For instance, based on the fatality rate of 33% (3/9), it would seem that bone and soft tissue cancer is associated with a high risk of COVID-19 fatality. However, the data contain only 9 patients with bone and soft tissue cancer; therefore, more information is needed to make clinical decisions for patients with COVID-19 and bone and soft tissue cancer. Clinicians can access additional information from the full text of papers from the Navigator to make better decisions about patient treatment.

Table 3. Outcome statistics for all tissue types.

Tissue type with cancer	Patients (n)	Deaths due to COVID-19 (% fatality), n (%)
Bladder/urinary tract	1	0 (0)
Bone and soft tissue	9	3 (33)
Bowel	5	0 (0)
Brain/central nervous system	2	0 (0)
Breast	91	5 (6)
Cervix	1	0 (0)
Genitourinary	4	0 (0)
Gynecological	180	20 (11)
Head and neck	8	0 (0)
Hematologic not specified	1893	492 (26)
Kidney	18	2 (11)
Liver	5	1 (20)
Lung	225	67 (30)
Lymphoid	504	152 (30)
Myeloid	29	5 (17)
Not specified	13265	2632 (20)
Ovary/fallopian tube	2	0 (0)
Pleura	1	0(0)
Prostate	124	28 (23)
Skin	3	0 (0)
Thoracic	200	66 (33)
Thymus	1	0 (0)

Discussion

Views From Clinician Users

We surveyed clinicians who tested the SMART COVID Navigator to elicit their assessment. Their feedback highlighted the following uses of the application in patient care and as an educational resource. The clinicians noted that the application allows for a real-time assessment of comorbidities that any given patient may have that could impact severity and fatality risk from COVID-19 and that it allows for the assessment of patients in which multiple factors may be at play. The clinicians also noted that the application could be used at the point of care, and this application filled an unmet clinical need. They stated that the application made it easier to narrow down the literature and find out the relevant patient management–related answers quickly. They also pointed out that the fatality and severity rates displayed by the CORD-19 data could help them in triaging and stratifying patients in limited resource conditions, allowing them to decide whether aggressive treatment was warranted immediately or not. This would also help them involve palliative care early enough if needed. Finally, they commented that the application would help them in shared decision-making with patients.

Apart from patient care, the clinicians felt that the SMART COVID Navigator could serve as an educational resource for teaching medical students, residents, journal clubs, and perhaps, even in continuing medical education for physicians. Some also felt that the application can help in making institutional guidelines. The clinicians made suggestions for future work related to the SMART COVID Navigator. Real-time use of the application would be enhanced by creating a smartphone interface in addition to the current desktop-based web interface, considering that physicians often perform literature reviews on smartphones. Although the application currently links to the VA and CMS, some clinicians felt that the application would benefit from being linked to other EHRs such as Epic. In fact, the same type of user interface and back-end API (SMART-on-FHIR) can be leveraged and is supported by these as well. Another suggestion was to link the application to a patient's COVID-19 vaccination records to further improve the physician's ability to treat the patient.

Additional Comments and Extensions

SMART COVID Navigator was created in response to a growing need for precision medicine tools to assist doctors dealing with COVID-19–infected patients with risk factors shown to affect COVID-19 progression. The Navigator achieves this by connecting patient medical information—in the form of EHRs—with data sets giving comprehensive information about patient outcomes. We show in this paper how to create a clinical decision support system for physicians to understand how a new disease (COVID-19) interacts with a broad set of patient risk factors through the use of the CORD-19 data set. In addition, we expand 1 existing patient condition, namely, cancer, and provide physicians information on how various types of cancers interact with COVID-19. This can be extended to other conditions such as heart disease since certain chronic conditions appear with a large number of variations in patients.

This application benefits from the creation of research study data sets such as CORD-19 and the Reboot: COVID-Cancer Project, thereby showing the value of such collaborative efforts to collect current research findings in one place. Currently, these data sets do not provide API access; therefore, we have to periodically download these data sets for use in the app. In the future, we recommend that such data sets provide API access to enable real-time updating. We note that in the CORD-19 data set, some disease categories are examined by a limited number of papers; thus, the results may not be meaningful for our understanding of COVID-19's impact on patients with that risk factor. Similarly, certain cancer tissue types are not well represented in the Reboot: COVID-Cancer Project data set. We expect that as research into COVID-19 progresses further, there will be better representation across disease categories and tissue types. The ability for users to log into 2 FHIR-supported EHR systems is a useful feature of the application. In addition to providing physicians the latest information about patient's health records, it can also alert users to discrepancies between different databases, giving them an opportunity to correct their health records. This application is not limited to just the 2 EHR databases currently used (VA and CMS); more EHRs can be added to the application as long as they follow the FHIR format, allowing for even more expansion of patient record access.

A possible next step for the SMART COVID Navigator could be to better consolidate the research available in CORD-19 and the Reboot: COVID-Cancer Project by using meta-analysis of the various studies to provide more information to the clinician [24]. For instance, more sophisticated weighting of the results in different papers could be implemented. Another extension would be to provide treatment information through the Navigator platform depending upon the patient's health profile. In addition, advancements in artificial intelligence and machine learning technologies could play an important role in the improvement of the Navigator platform, as well as precision medicine as a whole, by offering predictions about disease progression based on past patient health information, thereby tailoring treatment for the individual and allowing for increased efficiency in treatment [25]. This could take the form of creating patient risk scores generated based on data from EHRs.

Conclusion

Precision medicine has long been hailed as the next step in advancing patient care, with institutions such as the White House creating initiatives to further precision medicine technology advancement [26]. The SMART COVID Navigator is a clinical decision support tool designed to allow physicians to provide precision medicine in the context of COVID-19. Using the app, clinicians can identify patient risk factors from multiple EHRs and connect them to databases of COVID-19 research. Without a clinical decision support system supporting COVID-19 precision medicine, a clinician would need to examine multiple studies in order to fully understand disease progression and fatality outcomes, given a particular disease that their patient has, which is costly in terms of time and effort. The number of studies examining particular disease conditions and their relationship to COVID-19 is growing at a rapid pace. This application simplifies this by providing summary statistics on the risk factors' effect on COVID-19 severity and fatality.

Although this application currently focuses on COVID-19, it can be a readily available platform for quickly expanding into any potential new diseases that emerge.

Acknowledgments

The research reported in this publication was supported by grant National Cancer Institute U01 CA231840. Funding for SKB was provided in part by the Clinical and Translational Science Award (UL1 TR001439) from the National Center of Advancing Translational Sciences, National Institutes of Health, the University of Texas Medical Branch Claude D Pepper Older Americans Independence Center funded by National Institute on Aging, National Institutes of Health (P30 AG024832), and MD Anderson Cancer Center. RRM received research funding from Bayer, Pfizer, Tempus. JW received research funding from National Institutes of Health and American Association for Cancer Research related to this work.

Authors' Contributions

JW and GA contributed as cosenior authors.

Conflicts of Interest

RRM serves on Advisory Board for AstraZeneca, Aveo, Bayer, Bristol Myers Squibb, Calithera, Exelixis, Janssen, Merck, Novartis, Pfizer, Sanofi, and Tempus. RRM is a consultant for Dendreon, Myovant, Sorrento Therapeutics, and Vividion. RRM also serves on the molecular tumor board at Caris. None of these items are related to this work. JW is a consultant for Westat, Roche, and Flatiron Health. JW has ownership of HemOnc.org LLC, outside the submitted work.

References

1. What is precision medicine? MedlinePlus. URL: https://medlineplus.gov/genetics/understanding/precisionmedicine/definition/ [accessed 2021-07-06]
2. What is an electronic health record (EHR)? HealthIT.gov. URL: https://www.healthit.gov/faq/what-electronic-health-record-ehr [accessed 2021-07-06]
3. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. N Engl J Med 2020 Mar 26;382(13):1199-1207 [FREE Full text] [doi: 10.1056/NEJMoa2001316] [Medline: 31995857]
4. Wadman M. How does coronavirus kill? Clinicians trace a ferocious rampage through the body, from brain to toes. World Health Organization. 2020 Apr 18. URL: https://pesquisa.bvsalud.org/global-literature-on-novel-coronavirus-2019-ncov/resource/en/covidwho-71625 [accessed 2022-01-31]
5. Zaki N, Alashwal H, Ibrahim S. Association of hypertension, diabetes, stroke, cancer, kidney disease, and high-cholesterol with COVID-19 disease severity and fatality: A systematic review. Diabetes Metab Syndr 2020;14(5):1133-1142 [FREE Full text] [doi: 10.1016/j.dsx.2020.07.005] [Medline: 32663789]
6. McRae MP, Simmons GW, Christodoulides NJ, Lu Z, Kang SK, Fenyo D, et al. Clinical decision support tool and rapid point-of-care platform for determining disease severity in patients with COVID-19. Lab Chip 2020 Jun 21;20(12):2075-2085 [FREE Full text] [doi: 10.1039/d0lc00373e] [Medline: 32490853]
7. Wu G, Yang P, Xie Y, Woodruff HC, Rao X, Guiot J, et al. Development of a clinical decision support system for severity risk prediction and triage of COVID-19 patients at hospital admission: an international multicentre study. Eur Respir J 2020 Aug;56(2):2001104 [FREE Full text] [doi: 10.1183/13993003.01104-2020] [Medline: 32616597]
8. Warner JL, Prasad I, Bennett M, Arniella M, Beeghly-Fadiel A, Mandl KD, et al. SMART Cancer Navigator: A Framework for Implementing ASCO Workshop Recommendations to Enable Precision Cancer Medicine. JCO Precision Oncology 2018 Nov(2):1-14 [FREE Full text] [doi: 10.1200/po.17.00292]
9. Li X, Warner JL. A Review of Precision Oncology Knowledgebases for Determining the Clinical Actionability of Genetic Variants. Front Cell Dev Biol 2020;8:48 [FREE Full text] [doi: 10.3389/fcell.2020.00048] [Medline: 32117976]
10. Overby CL, Tarczy-Hornoch P. Personalized medicine: challenges and opportunities for translational bioinformatics. Per Med 2013 Jul 01;10(5):453-462 [FREE Full text] [doi: 10.2217/pme.13.30] [Medline: 24039624]
11. COVID application. GitHub. URL: https://github.com/smart-covid-navigator/Covid-Application [accessed 2022-01-31]
12. Li Z, Roberts K, Jiang X, Long Q. Distributed learning from multiple EHR databases: Contextual embedding models for medical events. J Biomed Inform 2019 Apr;92:103138 [FREE Full text] [doi: 10.1016/j.jbi.2019.103138] [Medline: 30825539]
13. Alterovitz G, Heale B, Jones J, Kreda D, Lin F, Liu L, et al. FHIR Genomics: enabling standardization for precision medicine use cases. NPJ Genom Med 2020;5:13 [FREE Full text] [doi: 10.1038/s41525-020-0115-6] [Medline: 32194985]
14. OAuth2. URL: https://oauth.net/2/ [accessed 2021-07-06]
15. OpenID Connect. URL: https://openid.net/connect/ [accessed 2021-07-06]
16. Kricka L, Polevikov S, Park JY, Fortina P, Bernardini S, Satchkov D, et al. Artificial Intelligence-Powered Search Tools and Resources in the Fight Against COVID-19. EJIFCC 2020 Jul;31(2):106-116 [FREE Full text] [Medline: 32549878]

17. Semantic Scholar. URL: https://www.semanticscholar.org/ [accessed 2021-07-06]
18. Lu WL, Lo K, Chandrasekhar Y, Reas R, Yang J, Eide D, et al. CORD-19: The COVID-19 Open Research Dataset. ArXiv. URL: https://arxiv.org/abs/2004.10706 [accessed 2021-07-06]
19. COVID-19 open research dataset challenge (CORD-19). Kaggle. URL: https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge [accessed 2021-07-06]
20. SNOMED International. URL: http://www.snomed.org/ [accessed 2021-07-06]
21. HL7 FHIR: ICD-10. URL: http://hl7.org/fhir/sid/icd-10 [accessed 2021-07-06]
22. Kuderer N, Choueiri TK, Shah D, Shyr Y, Rubinstein S, Rivera D, et al. Clinical impact of COVID-19 on patients with cancer (CCC19): a cohort study. The Lancet 2020 Jun;395(10241):1907-1918 [FREE Full text] [doi: 10.1016/S0140-6736(20)31187-9]
23. Reboot: COVID-cancer project. Reboot Rx, Inc. URL: https://rebootrx.org/covid-cancer [accessed 2021-07-06]
24. Zaki N, Mohamed E, Ibrahim S, Khan G. The influence of comorbidity on the severity of COVID-19 disease: A systematic review and analysis. medRxiv. 2020 Aug 11. URL: https://www.medrxiv.org/content/10.1101/2020.06.18.20134478v2 [accessed 2020-08-16]
25. Schork N. Artificial Intelligence and Personalized Medicine. Cancer Treat Res 2019;178:265-283 [FREE Full text] [doi: 10.1007/978-3-030-16391-4_11] [Medline: 31209850]
26. Collins F, Varmus H. A New Initiative on Precision Medicine. N Engl J Med 2015 Feb 26;372(9):793-795 [FREE Full text] [doi: 10.1056/nejmp1500523]

Abbreviations

API: application programming interface
CMS: Center for Medicare and Medicaid Services
CORD-19: COVID-19 Open Research Dataset
EHR: electronic health record
FHIR: Fast Healthcare Interoperability Resources
ICD-10: International Classification of Diseases-tenth version
VA: Veterans Affairs

Edited by C Basch; submitted 02.04.21; peer-reviewed by J Walsh, J Florez-Arango; comments to author 17.05.21; revised version received 07.09.21; accepted 07.10.21; published 18.02.22.

Please cite as:
Suraj V, Del Vecchio Fitz C, Kleiman LB, Bhavnani SK, Jani C, Shah S, McKay RR, Warner J, Alterovitz G
SMART COVID Navigator, a Clinical Decision Support Tool for COVID-19 Treatment: Design and Development Study
J Med Internet Res 2022;24(2):e29279
URL: https://www.jmir.org/2022/2/e29279
doi:10.2196/29279
PMID:34932493

Original Paper

The Effects of Website Traits and Medical Skepticism on Patients' Willingness to Follow Web-Based Medical Advice: Web-Based Experiment

Jennifer Claggett[1], PhD; Brent Kitchens[2], PhD; Maria Paino[3], PhD; Kaitlyn Beisecker Levin[4], MD

[1]School of Business, Wake Forest University, Winston-Salem, NC, United States

[2]McIntire School of Commerce, University of Virginia, Charlottesville, VA, United States

[3]Department of Sociology, Anthropology, Social Work, and Criminal Justice, Oakland University, Rochester, MI, United States

[4]WholeHealth Medical, Charlottesville, VA, United States

Corresponding Author:
Jennifer Claggett, PhD
School of Business
Wake Forest University
1834 Wake Forest Rd
Farrel Hall
Winston-Salem, NC, 27109
United States
Phone: 1 1336302799
Email: claggett@gmail.com

Abstract

Background: As people increasingly turn to web-based sources for medical information, we offer some insight into what website traits influence patients' credibility assessment. Specifically, we control for brand and content length, while manipulating three website traits: authorship, format, and tone. Furthermore, we focus on medical skepticism to understand how patients with high levels of medical skepticism may react to web-based medical information differently. Medical skepticism is related to a patient's doubts about the value of conventional medical care; therefore, skeptics may have different practices and criteria when conducting their own web-based medical searches.

Objective: The aim of this study is to evaluate how website traits affect the likelihood that patients follow web-based medical advice and how this varies among patients with differing levels of medical skepticism.

Methods: This web-based experiment presented participants with a hypothetical medical situation about leg cramps and offered a website with treatment advice. We varied the websites the participants observed across three traits: authorship (patient or physician), format (article or discussion forum), and tone (objective or experience-based). The 2305 participants were randomly assigned to 1 of 8 possible conditions and then asked the extent to which they would follow the advice. Health care patterns and coverage, demographics, and the participants' level of medical skepticism were captured.

Results: Our participants were selected to be demographically representative of the population of internet users in the United States. The 2305 complete responses were analyzed with ordinary least squares regression. Our analysis reveals that people are more likely to accept web-based medical advice authored by a physician (P<.001) and presented with an objective tone (P=.006), but these preferences erode as the levels of medical skepticism increase. Medical skepticism was measured by means of a previously established index on a 0 to 4 scale, and the average score was 2.26 (SD 0.84). Individuals with higher levels of medical skepticism were more likely to follow web-based medical advice in our experiment (P<.001). Individuals with low levels of medical skepticism found the discussion forum format more credible, whereas those with high levels of medical skepticism preferred the article format (P=.03). We discuss the interactions between medical skepticism and all 3 website traits manipulated in the experiment.

Conclusions: Our findings suggest that, generally, physician authorship and an objective tone create more persuasive web-based medical advice. However, there are differences in how patients with high levels of medical skepticism react to web-based medical resources. Medical skeptics are less discerning regarding the author's credentials and the presentation tone of the information. Furthermore, patients with higher levels of medical skepticism prefer article format presentations, whereas those with lower levels of medical skepticism prefer discussion forum–style formatting.

(J Med Internet Res 2022;24(2):e29275) doi:10.2196/29275

KEYWORDS

web-based information credibility assessment; website traits; medical skepticism; mobile phone

Introduction

Background

Skepticism toward medical care and treatment has been a long-standing issue in health care [1], ranging from the development of antivaccination beliefs [2,3] and affecting how smokers perceive health-risk information [4] to affecting which treatments people with arthritis decide to use [5]. Medical skepticism, specifically, is related to a patient's doubts about the value of medical care and is defined as "global doubts regarding the ability of conventional medical care to appreciably alter health status" and has been shown to correlate with higher rates in mortality at follow-up [6]. However, other studies have noted the potential benefits of medical skepticism, such as medically skeptical older adults reporting higher levels of self-rated overall health [7]. Previous work has focused on which demographics are more inclined to be skeptical of health care, generally concluding that it is highest among young White people with less education and often lacking health insurance [8,9]. Even so, medical skepticism is widespread, and the trend may be rooted in a consumerist movement where patients are encouraged to be more involved in their health care, making it of particular interest to understand how their decision-making process regarding medical care unfolds [10].

The internet is an obvious tool for patients to find alternative medical information [11,12]. All patients, regardless of their medical skepticism, increasingly use web-based resources to seek health information [13,14]. A Pew Research Center report found that 80% of US internet users (>93 million people) have searched for health-related topics on the web, and most of the health-related searches are related to specific medical problems they are experiencing [15]. In 1 survey, 80% of the physicians confirmed experiences of patients presenting internet-sourced medical information [16]; and in another survey, 41% of the patients stated that they had used a medical professional to confirm their own internet diagnoses [15]. However, this growing interest in web-based medical information has been answered by an explosion of available health content, ranging from predatory websites that spread misinformation and nonprofit institutions creating web-based article repositories to patients using interaction-enabled web-based formats to discuss medical information in a question-and-answer format [17-19]. As gathering health information on the web has been established as a norm, it is important that we understand what causes a patient to follow or ignore medical advice they read on the web [20-24]. Logically, medically skeptical patients may also have different web search behaviors and react to website traits differently. Therefore, we consider what web-based content traits drive people's willingness to follow the advice, as well as the role medical skepticism plays in the web-based credibility assessment process.

Previous research has looked at the use of patient portals as a web-based form of communication between patients and physicians [25-27], but portals provide a web-based communication mechanism that supplements a previously established relationship between a physician and a patient. Prior research has also considered how active participation in web-based health care communities may influence patient outcomes and provide emotional support [28-31]; yet, this assumes that a patient has joined, and is actively participating in, a community, usually sharing their own details and questions. In contrast, our study aims to better understand how patients decide whether to follow web-based medical advice in more casual web research scenarios, specifically those that do not necessarily require the intervention of a physician or long-term emotional support. To that end, we focus on three potential traits that shape web content and may affect credibility assessment: authorship, format, and tone of presentation.

The objectives of this study are to (1) determine what website traits affect a patient's likelihood to follow the presented medical advice and (2) determine whether patients with higher levels of medical skepticism exhibit different patterns of website traits influencing their likelihood to follow medical advice found on the web. Our findings increase our understanding of how patients decide whether to follow web-based medical advice and may inform the design of health care websites. Furthermore, we extend our knowledge about medical skepticism by finding evidence that it alters the credibility assessment that patients undertake when considering web-based medical advice.

Website Traits

Overview

There has been an explosion in medical websites in the last several decades that individuals may turn to for obtaining advice, sharing experiences, voicing concerns, and informing decisions [32-34]. In any context, web users must look for cues signaling that the content is credible and trustworthy [35,36]. We identify three specific website traits that vary across common health information websites: authorship, format, and presentation tone.

Authorship

The authorship of content is a major factor in how people assess the credibility of information [37]. A meta-analysis of studies looking at the effects of health care expertise on the credibility of health information suggests that authorship of web-based content is important and experts are favored [38]. However, the research is not consistent, and in some situations, patients prefer nonexpert advice [39] or perceive a source as credible because of positive judgments about the trustworthiness of the author instead of professional credentials [21]. For simplicity, we focused on two major categories of authors who often provide web-based health care information: physicians and other patients. Physicians represent the easy-to-recognize role of a health care professional who has formal training and expertise, whereas other patients are peers who may have faced similar medical situations.

Format

The increasing prevalence of user-generated content is one of the strongest trends of the past decade [40,41]. This has created new types of web-based content fundamentally different from the more traditional static webpage format, especially regarding navigation design and connectedness [42]. For example, a medical article is usually presented as a static page and is read-only. The owner of a static webpage is also the controller of the information presented. For brevity, we refer to this static, read-only delivery as article format. In contrast, web forums or blogs invite users to also create and share content (eg, ask questions or respond to previous posts) in a dynamic and participative manner [41]. This cocreation process means that the webpage owner is not solely responsible for generation of the content that is visible. We refer to this dynamically cocreated content as discussion forum format. The article format presents a single point of view in a controlled environment, whereas the discussion forum format often presents a cocreated set of advice.

Tone of Presentation

Medical information can be presented on the web as objective or experiential [9]. Objective information is content presented as fact, devoid of any personal attachment or interpretation by the person conveying the information. Experiential information is presented as derived from the actual experiences and insights of the person conveying the information. Some previous research makes the assumption that information provided by health care professionals is objective and information provided by other patients is experiential [28]; however, we believe that the 2 should not be conflated but instead considered separately. Health care professionals can offer advice because of their experiences with other patients, and other patients can convey information as objective fact. Broadly speaking, people tend to appreciate experiential information in decision-making contexts because it signals familiarity with the content [21]; yet, traditionally, an objective presentation signals that the information is well established and broadly accepted [43].

Methods

Sampling and Data Collection

We conducted a web-based experiment through a web-based survey. Qualtrics software was used to recruit participants, and after excluding those with missing or incomplete data, we obtained a sample of 2305 participants. Panelists were recruited to be demographically representative of the population of internet users in the United States, which is a strength of Qualtrics panels [44]. We excluded participants who worked in a medical profession and those who had previous experience with leg cramps (the focus of our experimental manipulation) to remove participants with previous expertise in the study's context. We captured demographics (sex, race, education level, income, age, and geography: whether they lived in a rural, suburban, or urban area), health care situation (number of recent health care visits, current method of receiving primary care, and health insurance status), and the respondent's level of medical skepticism.

Each participant was presented with the same hypothetical medical situation to begin the experiment: "Your friend has recently been battling leg cramps. They ask you to help them research the condition online, and you find the following resource. Please read the web page presented and then answer the questions related to your experience with the online resource."

This situation was chosen such that there would be a nontrivial treatment recommendation that could nonetheless be administered without formally seeking professional medical intervention. Next, 1 of 8 different webpages containing health information regarding leg cramps was presented to the participant. The set of 8 webpages used in the experiment (2×2×2 experimental design) represents each possible permutation of the three website traits of interest: authorship (physician or patient), format (article or discussion forum), and tone (objective or experiential). Each prompt provides identical treatment advice regarding leg cramps, regardless of the website traits shown in the experimental manipulation to which they were assigned.

A positive perception of a specific brand or the reputation of a website increases the likelihood that people will follow the presented advice [42,45-47]. Opinions of friends and acquaintances about a particular website source might also influence patients because they build brand recognition [48,49]. In addition, the length of the text and other formatting features that affect the ease of skimming (eg, bullet points vs long blocks of text) are known to affect the likelihood that people will read the advice and then follow that advice [50-52]. To avoid confounding effects rooted in brand recognition, text length, or readability, our experimental design eliminates these factors by removing branding and standardizing the text length and grammatical presentation of the content shown to participants across all 8 website prompts (including order of information and sentence structure to the extent possible). Example prompts are presented in Figures 1 and 2; all 8 are available upon request.

Individuals were randomly assigned to each experimental condition, and there were between 269 and 301 individuals in each of the 8 possible website version groups. After viewing the prompt, participants were asked to indicate the likelihood that they would recommend the presented advice to their friend with leg cramps.

Figure 1. Experimental website prompt displaying physician authorship, article format, and experiential presentation tone.

How to Stop Leg Muscle Cramps

M Weatherspoon, M.D.
July 6, 2018

What's going on?

We experience muscle cramps when a muscle involuntarily contracts on its own. I had a patient last week describe her muscle cramp as a hard lump at the point of pain — that's the contracted muscle. Cramps usually occur for a reason. My patient hadn't strained a muscle, so her cramping was probably because of muscle fatigue or overuse. Her body may have been dehydrated, or she wasn't getting enough electrolytes, such as potassium or magnesium. Those minerals help our muscles work more smoothly, and fluids help our body process the minerals.

Most cases of muscle cramps do not indicate a worrisome underlying condition. Cramps might be related to alcoholism, hypothyroidism, or diabetes. If the frequency of your cramps bothers you, tell your doctor.

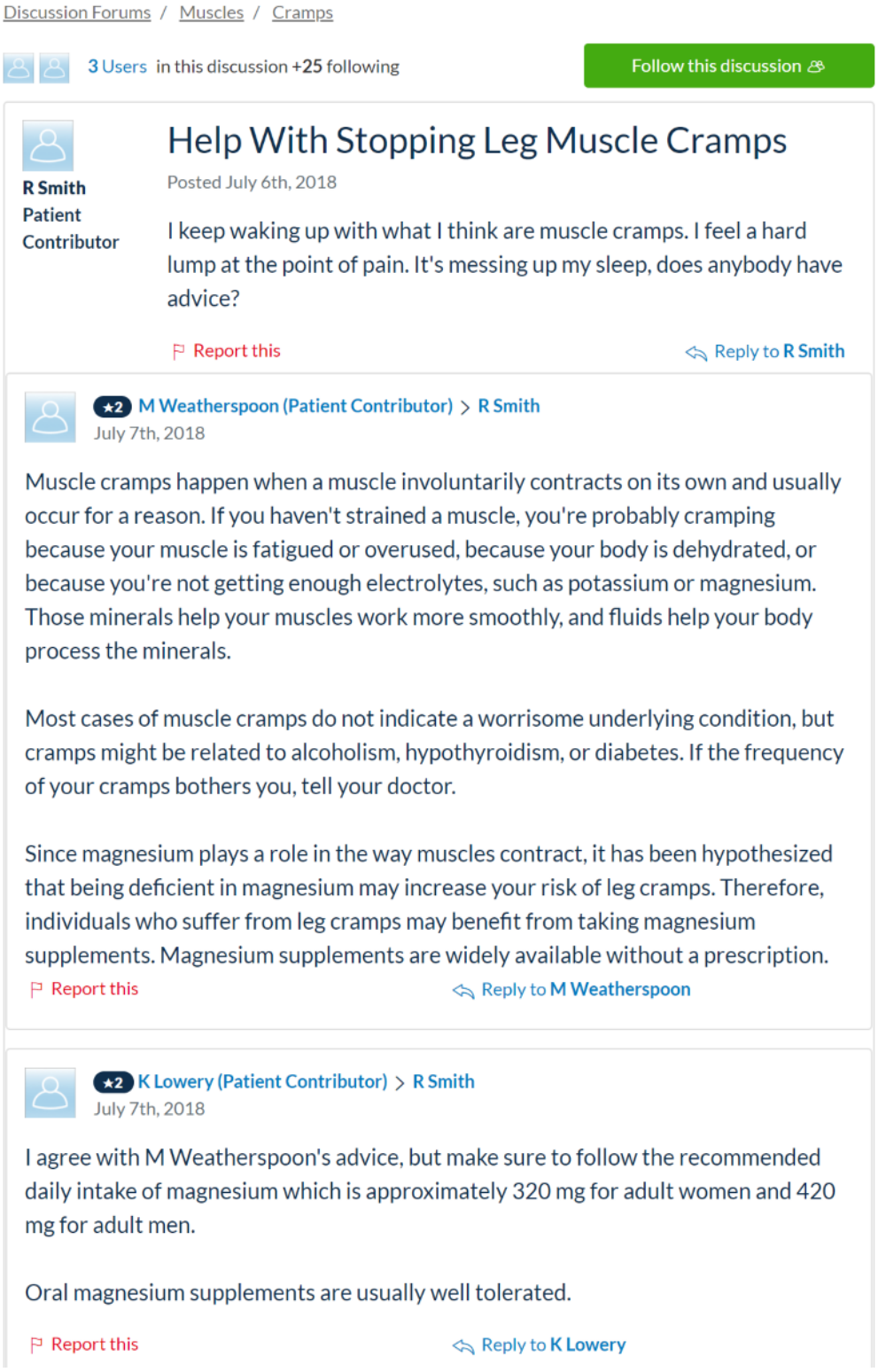
Magnesium Supplements

Since magnesium plays a role in the way muscles contract, it has been hypothesized that being deficient in magnesium may increase your risk of leg cramps. Therefore, I advised my patient that taking magnesium supplements may be beneficial. Magnesium supplements are widely available without a prescription.

Additional Considerations

I recommend taking the daily intake of magnesium which is approximately 320 mg for adult women and 420 mg for adult men. Oral magnesium supplements are usually well tolerated by my patients.

Figure 2. Experimental website prompt displaying patient authorship, discussion forum format, and objective presentation tone.



Outcomes and Measures

Follow Web-Based Medical Advice

Our dependent variable was a continuous variable, ranging from 0 to 4, where lower numbers indicated that individuals were less likely to follow web-based medical advice and higher numbers indicated that individuals were more likely to follow web-based medical advice. This variable was an index composed of four ordinal measures indicating the extent to which individuals agreed with the following: (1) recommending that their friends try magnesium supplements, (2) the extent to which individuals believe that trying magnesium supplements is good advice, (3) if an individual had leg cramps, the extent to which they agree that they would try magnesium supplements, and (4) the webpage convinced them that they should try magnesium supplements to reduce leg cramps.

Individuals reported the extent to which they agreed with each of these statements by responding to a Likert scale that ranged

from *strongly disagree* to *strongly agree*. We averaged these measures to form a continuous index for following web-based medical advice (Cronbach α=.94).

Independent Variables

Experimental Conditions

We included three independent variables to account for each of the aforementioned experimental conditions: the author (physician or patient), the format (article or discussion forum), and the tone (objective or experiential).

Medical Skepticism

We used an index to measure medical skepticism, and this index was based on 3 characteristics that captured an individual's skepticism toward medicine. Medical skepticism was measured by means of a previously validated scale [8,53]. Respondents were asked the extent to which they agreed (on a scale of 0-4) with the following questions: (1) I can overcome most illness without help from a medical professional, (2) home remedies are better than medicines prescribed by doctors, and (3) I understand my health better than most physicians do (Cronbach α=.62). Although the reliability score is below the recommended threshold of Cronbach α=.70, it is an established scale [8,53] and similar to the reliability score reported in previous studies (eg, Cronbach α=.69 in the study by Fiscella et al [6] and Cronbach α=.63 in the study by Jensen et al [10]).

Health Care Situations

We used several independent variables to account for individuals' health care situations. These were important aspects to capture for both our research objectives. A patient's medical care situation and exposure may influence how they interpret web-based health care information [54,55], in general, and previous work has noted that medical skeptics may seek out in-person health care differently [8]. To accurately measure and control for these variations, we first included the approximate number of health care visits that respondents had made in the last 2 years because of illness or injury. Health care visits included visits that respondents have made for themselves or as a companion to a friend or family member because either visit would expose the individual to conversations with individual health care providers. Second, we used a measure that indicated the location where respondents sought primary medical care. This was a dichotomous measure that showed whether a respondent used emergency or urgent care as their primary care or whether they used a clinic, primary care physician, or other health care facility or whether the respondent did not seek any medical care. Finally, we included a variable that measured whether the respondent had health insurance. We compared those who did not have health insurance with individuals who had any type of health insurance (ie, Medicaid, Medicare, employer-provided, military or veteran, self-funded).

Control Variables

We included geographical setting and demographic data in our models, including sex, race, age, income, and educational attainment (Table 1).

Table 1. Descriptive statistics (N=2305).

Variable	Description	Value
Dependent variable: follow web-based medical advice, mean (SD)	This is a continuous variable, ranging from 0 to 4, indicating the likelihood that the respondent will follow the web-based medical advice provided at the end of the prompt (Cronbach α=.94)	2.97 (0.90)
Experimental conditions: 2 × 2 × 2 design involving the following three factors, n (%)		
Physician	Dichotomous variable indicating that the experimental prompt was written by a physician (as opposed to a patient)	1170 (50.76)
Article	Dichotomous variable indicating that the experimental prompt was written in article format (as opposed to a discussion forum format)	1171 (50.8)
Objective	Dichotomous variable indicating that the experimental prompt was written in an objective manner (as opposed to experiential manner)	1158 (50.24)
Medical skepticism, mean (SD)	This is a continuous variable, ranging from 0 to 4, indicating the extent to which the respondent is skeptical toward medicine (Cronbach α=.62)	2.26 (0.84)
Health care situation		
Number of health visits, mean (SD)	The approximate number of health care visits in the last 2 years (as a patient or accompanying a friend or family member)	4.40 (3.47)
Primary care is emergency or urgent, n (%)	Dichotomous variable indicating that the respondent uses emergency or urgent care as their primary or usual care option (as opposed to other sources such as primary care physician or a community clinic)	477 (20.69)
No insurance, n (%)	Dichotomous variable indicating that the respondent has health insurance (as opposed to being without insurance)	344 (14.92)
Home geographical category, n (%)		
Rural	Respondent lives in a rural setting	558 (24.21)
Suburban	Respondent lives in a suburban setting	1046 (45.38)
Urban	Respondent lives in an urban setting. This is the reference category	701 (30.41)
Demographics		
Female, n (%)	Respondent is female (the reference category is male)	1352 (58.66)
Asian or Pacific Islander, n (%)	Respondent identifies as Asian or Pacific Islander	158 (6.85)
Black, n (%)	Respondent identifies as Black	511 (22.17)
Hispanic or Latino, n (%)	Respondent identifies as Hispanic or Latino	210 (9.11)
Other race, n (%)	Respondent identifies as other race	88 (3.82)
White, n (%)	Respondent identifies as White. This is the reference category	1338 (58.05)
Bachelor's degree, n (%)	Dichotomous variable indicating that the respondent has at least a bachelor's degree (reference category is less than a bachelor's degree)	689 (29.89)
Income, mean (SD)	An ordinal variable, ranging from 0 to 11, entered into the model as a continuous predictor because of its underlying interval–ratio nature. The answer 4 (the average) indicates a salary of approximately US $30,000 to US $39,999	4.01 (3.16)
Young, n (%)	Respondent is young: aged 18-34 years	728 (31.58)
Middle-aged, n (%)	Respondent is middle-aged: aged 35-64 years	1223 (53.06)
Older, n (%)	Respondent is older: aged 65 to ≥85. This is the reference category	354 (15.36)

Data Analysis

We used an ordinary least squares (OLS) regression model to analyze how factors such as the experimental condition, medical skepticism, and demographic characteristics contributed to the extent to which individuals were willing to follow the advice offered for a leg cramp medical condition. Our dependent variable was continuous in nature, thus making OLS regression the most appropriate analytical model.

Ethics Approval

This study was approved by the University of Virginia's Institutional Review Board for Social and Behavioral Science (Project Review 2018-0398-00).

Results

Participant Characteristics

Of the 2305 participants in our study, 1352 (58.66%) were women. Our sample was racially diverse, with 58.05% (1338/2305) of the participants identifying as White, 22.17% (511/2305) as Black, 9.11% (210/2305) as Hispanic or Latino, 6.85% (158/2305) as Asian or Pacific Islander, and 3.82% (88/2305) as a different race. Of the 2305 participants, 1223 (53.06%) were aged 35-64 years and 728 (31.58%) were aged 18-34 years, whereas 354 (15.36%) were aged >65 years.

On average, the respondents reported approximately 4.40 (SD 3.47) health care visits in the last 2 years (to accompany a friend or family member or as patients themselves), and 20.69% (477/2305) of the respondents indicated that their primary care occurred in an emergency room or urgent care clinic. Most of our respondents carried some sort of insurance (eg, Medicare, Medicaid, and private insurance), but 14.92% (344/2305) reported having no insurance provider. Our respondents included those who lived in suburban (1046/2035, 45.38%), urban (701/2305, 30.41%), and rural (558/2305, 24.21%) areas. Of the 2305 respondents, 689 (29.89%) had a bachelor's degree or higher, and the average respondent reported an income generally consistent with that of the middle class [56] (Table 1).

Following Web-Based Medical Advice

At the end of each prompt, the respondents were provided with the same treatment advice regarding leg cramps, regardless of the website traits shown in the experimental manipulation. We present OLS regression results (Table 2) that assess how experimental conditions and medical skepticism contribute to the likelihood that respondents will follow this advice. In model 1 (Table 2), we focus on the direct relationships between our independent variables of interest and the likelihood that respondents follow web-based medical advice. Our findings suggest that website traits do affect whether individuals are likely to follow medical advice. When presented with an example written by a physician, the respondents were significantly more likely to follow advice (P<.001). Similarly, an objective writing style, as opposed to a tone reflecting personal experience, was positively associated with following web-based medical advice (P=.006).

Table 2. Ordinary least squares regression analysis of following web-based medical advice regressed on experimental conditions.

	Model 1	*P* value	Model 2	*P* value
Experimental conditions				
Physician author	0.184	<.001	0.385	<.001
Article format	0.019	.60	–0.196	.06
Objective writing style	0.101	.006	0.312	.003
Medical skepticism	0.099	<.001	0.144	.001
Interactions				
Skepticism × physician	__[a]	—	–0.089	.04
Skepticism × article	—	—	0.094	.03
Skepticism × objective	—	—	–0.094	.03
Health care				
Number of health care visits	0.025	<.001	0.025	<.001
Primary care is emergency or urgent	0.099	.04	0.094	.047
No insurance (reference: any insurance)	–0.018	.73	–0.012	.83
Control variables				
Setting (reference: urban setting)				
Rural	–0.072	.18	–0.074	.16
Suburban	–0.009	.84	–0.009	.84
Female	0.010	.80	0.013	.73
Race (reference: White)				
Asian	–0.114	.14	–0.121	.12
Black	–0.042	.40	–0.045	.36
Hispanic	0.015	.83	0.019	.78
Other race	–0.075	.44	–0.079	.42
Bachelor’s degree or higher (reference=less than bachelor’s degree)	–0.232	<.001	–0.225	<.001
Income	0.007	.28	0.007	.31
Age (years; reference: older adults aged ≥65 years)				
Young (18-34)	–0.244	<.001	–0.242	<.001
Middle-aged (35-64)	–0.121	.03	–0.120	.03
Constant	2.678	<.001	2.576	<.001
R^2 (N=2305)	0.059	N/A[b]	0.060	N/A

[a]Not included in base model.

[b]N/A: not applicable.

As medical skepticism increased, individuals were more likely to follow web-based medical advice (*P*<.001). This may be because these individuals are more receptive to self-service sources of information as a substitute for the advice of health care providers, whom they distrust. Individuals who reported more frequent health care visits in the last 2 years were also positively associated with following web-based medical advice (*P*<.001). When individuals used the emergency room or urgent care as their primary source of health care, they were more likely to follow the presented advice (*P*=.04). The sex, race, and income of our respondents were not significantly related to whether they followed web-based medical advice, but education and age were significant predictors. Compared with those without higher education, individuals who had attained at least a bachelor’s degree were significantly less likely to follow web-based medical advice (*P*<.001). In general, younger individuals were less likely to follow web-based medical advice (*P*<.001 for young adults and *P*=.03 for middle-aged adults), which is consistent with previous research that found that younger patients were less likely to believe that providers listened to them [57] and less likely to seek medical care [58].

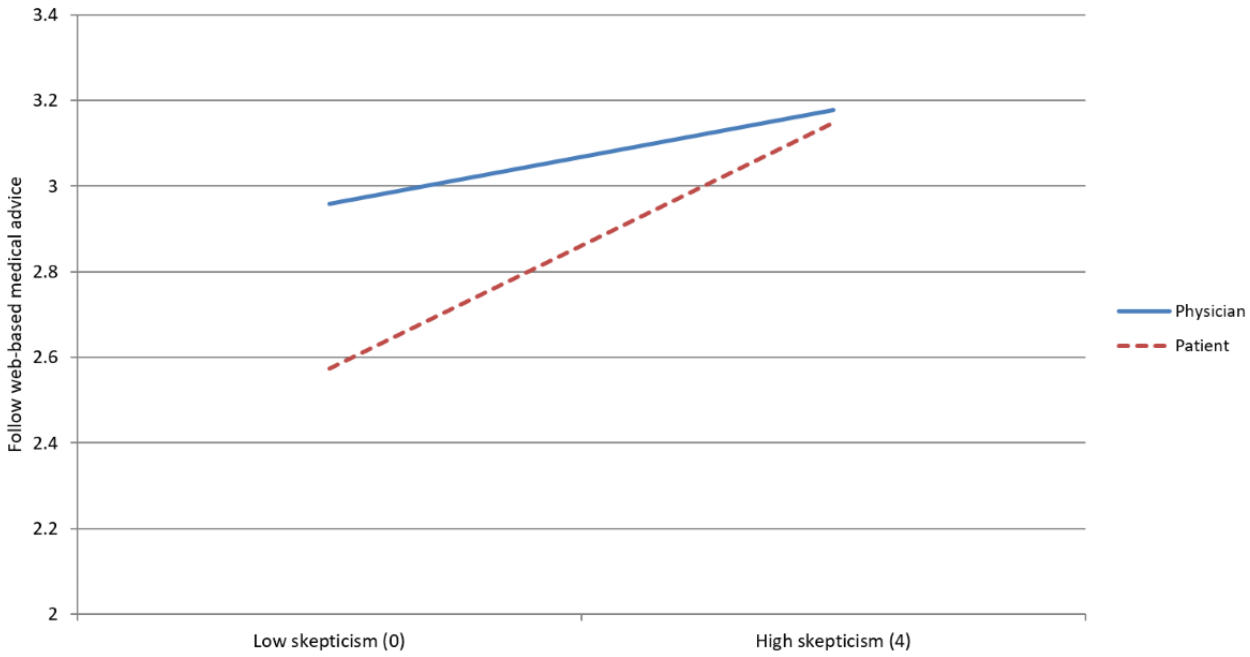
Interaction Between Medical Skepticism and Website Trait Influence

In model 2 (Table 2), we additionally examine moderation effects that capture the complex relationship between experimental conditions and medical skepticism. Model 1 demonstrates that individuals reporting higher levels of skepticism were more likely to follow web-based medical advice, regardless of the experimental condition. However, the significant coefficients for our 3 interaction terms indicate that the association between experimental condition and following web-based medical advice does vary by the degree of medical skepticism.

Physician Authorship × Medical Skepticism

Model 1 shows that physician advice (as opposed to patient advice) was more likely to be followed. However, as model 2 and Figure 3 show, this effect was moderated by medical skepticism such that the physician advantage was much smaller among respondents with higher skepticism compared with those with lower skepticism (*P*=.04). Follow-up regression models indicated that the physician authorship advantage was nonsignificant among respondents in the highest tertile of skepticism, but it was relatively large and statistically significant among those in the lowest tertile of skepticism (b=0.202; *P*=.05). Despite this erosion of physician authorship advantage among more skeptical users, individuals across all skepticism levels favored physician-authored web-based material—in no case were respondents more likely to follow web-based medical advice in the experiment after reading a webpage with a patient author.

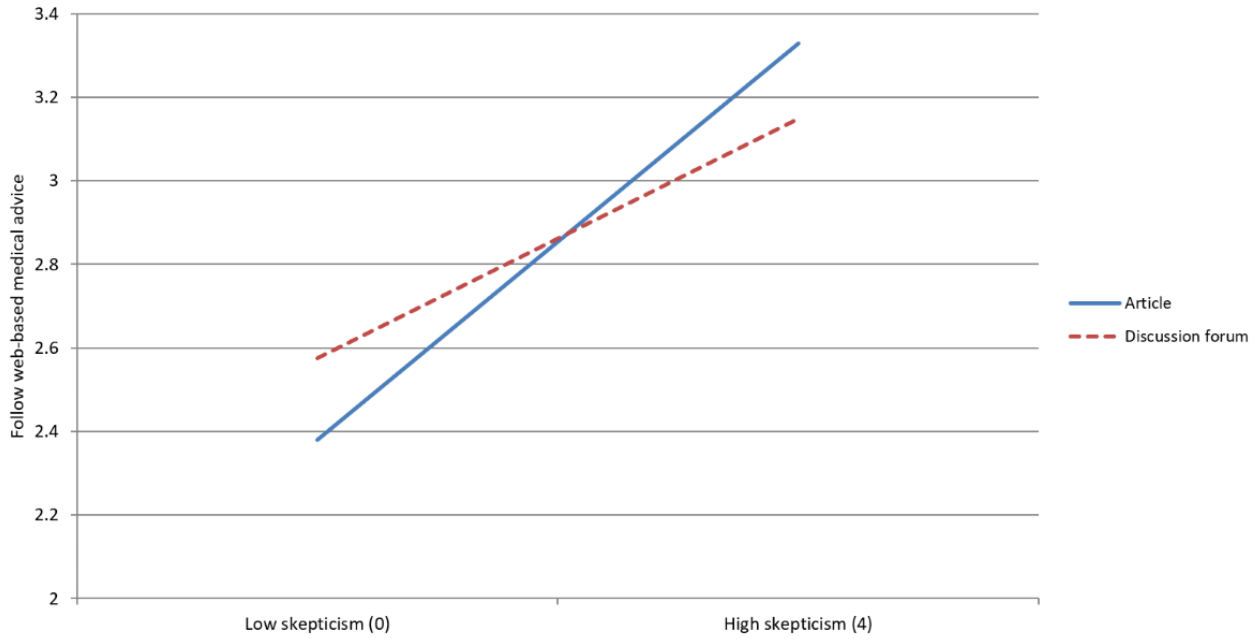
Figure 3. Interaction between medical skepticism and authorship.



Webpage Format × Medical Skepticism

In model 1, the format of the webpage (article vs discussion forum) was not significantly related to whether individuals followed web-based medical advice. However, model 2 and Figure 4 demonstrate that the effect of a webpage format was moderated by medical skepticism (*P*=.03). Those with lower levels of skepticism were more likely to follow web-based medical advice when presented with a discussion forum format. Follow-up regression models indicated a statistically significant discussion forum (vs article) advantage for respondents in the lowest tertile of medical skepticism (b=–0.270; *P*=.008), whereas those in the highest tertile of skepticism significantly preferred an article format (b=0.270; *P*=.008). This suggests that individuals with high levels of medical skepticism are swayed by an article format, but those with low levels of medical skepticism show a preference for a discussion forum format.

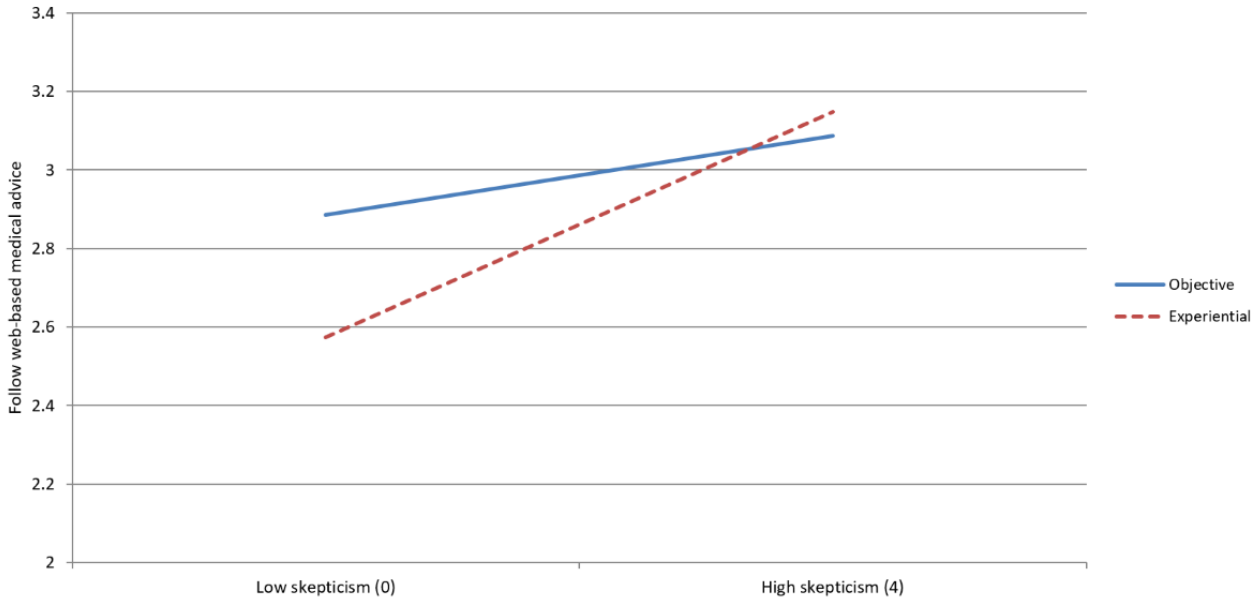
Figure 4. Interaction between medical skepticism and format.



Content Tone × Medical Skepticism

Our final interaction term illustrates how individuals with the highest level of medical skepticism have diminished content tone preferences. Model 2 and Figure 5 show how the effect of the content tone was moderated by medical skepticism (*P*=.03). Follow-up regression models demonstrate that the objective tone advantage was statistically significant among respondents in the lowest tertile of skepticism (b=0.240; *P*=.02). On average, people are more likely to follow objectively written advice, but that advantage disappears among those with higher skepticism.

Figure 5. Interaction between medical skepticism and writing style.



Discussion

Principal Findings

Our results demonstrate that, within our experimental condition of leg cramp advice, website presentation traits of medical information on the web do matter. Generally, patients are more receptive to medical information authored by a physician and presented in an objective tone, with this advantage diminishing among those with higher skepticism. The website format trait (article format vs discussion forum format) is more complex. Although our findings in model 1 suggest this trait is not a significant influencer when considering all participants on average, Figure 4 helps explain that this is due to different preferences depending on the level of medical skepticism, with low-skepticism participants preferring the discussion forum format and high-skepticism participants preferring the article format. This distinction implies that different website formats may be better suited for audiences of different levels of medical skepticism.

Medical skepticism has been previously identified as an important trait to consider when trying to understand patient behavior [5,6,8,9,59]. Learning about medical conditions and taking a more active role in health decisions may manifest in higher medical self-efficacy and simultaneously make patients more skeptical about physician advice [59]. For example, high levels of medical skepticism were associated with the use of complementary and alternative medical treatments in patients with arthritis [5]. We extend our understanding of medical skepticism and show evidence that it is also important in how people consume web-based medical information. We find that people consume and judge web-based medical advice credibility differently, depending on their level of medical skepticism. Patients with a greater degree of medical skepticism are more likely to follow web-based medical advice, regardless of the website traits. This insight makes sense because they are likely more familiar with getting medical advice or ideas from web-based sources because they conduct their own research instead of seeking out and listening to medical professionals in person. This underscores the ever-deepening power of web-based medical advice on patient behaviors, often at the expense of seeking in-person medical diagnoses.

Furthermore, the preferences of physician authorship over patient authorship and an objective writing tone over an experiential writing tone diminish as medical skepticism increases. This implies that medical skeptics may be more open to a wider variety of web-based medical advice. They are receptive to patient and physician accounts as well as experiential recounts, in addition to objective information. However, patients with a greater degree of medical skepticism find the article format more credible than the discussion forum format, which is opposite of the preferences of those with lower degrees of medical skepticism. This is an interesting finding, and may be due to a sense of legitimacy conveyed in an article format to a population more interested in researching and making their own medical decisions.

This research suggests that medical organizations should consider website traits when designing solutions to communicate medical advice. Physician-authored information written in objective tones in a discussion forum format seems to be a potentially effective combination for individuals with low medical skepticism. We do not formally survey web-based information sources to determine the relative frequency of presentation formats, but this combination seems to be rare, presenting a potentially significant opportunity. Our findings may also be useful to health care professionals because they interact with patients who are likely bringing their own ideas and information gathered from web sources into consultations. Those with higher levels of medical skepticism tend to be less discriminating about website traits, although they find web-based articles more persuasive.

Limitations and Future Research Directions

The hypothetical situation of a leg cramp scenario might have influenced participants in certain ways that limit the broader generalizability of our insights. The strength of an experiment is the ability to control factors in the participant's situation, but the experimental scenario may not accurately predict how real-world scenarios unfold [60]. Care was taken to select a medical scenario that (1) was plausible to research independently (ie, not so severe as to need immediate medical attention) and (2) could involve medical treatment advice that was not trivial but did not necessitate physician oversight. We also screened out participants who had previous experience with leg cramps in an attempt to recruit unbiased participants without existing opinions about leg cramp treatments.

Patients read internet content on a variety of devices, including desktop computers, laptops, tablets, and smartphones [61]. Our data reflect this variety, with 54.49% (1259/2305) of the participants accessing the study from mobile devices. Therefore, we believe that our findings are generalizable to a variety of consumer devices used by patients. However, future work would benefit from studying user device preferences more directly to see whether they alter patient acceptance of web-based medical advice.

Our website format distinction was between an article format (static website that the owner controls) and a discussion forum format (dynamic cocreation among multiple users). It is worth noting that there is a plethora of discussion forum styles and blogs, with their own different sets of features and uses. Our design incorporated a simple discussion forum layout, but future research could extend this research question to determine which set of traits and functions in a discussion forum layout distinctly affect patients' willingness to follow web-based medical advice.

Our study focused on nonchronic ailments for which patients may seek web-based medical advice. Patients who have chronic ailments may exhibit different behaviors regarding their assessment and adherence to web-based medical advice. For example, previous work has examined the importance of web-based communities for patients with chronic diseases [62-64]. Another extension of this research would be to consider a broader decision-making process of the patient. Our experimental design focused only on the participant's initial assessment of a single isolated webpage. In reality, patients may look at multiple webpages to form opinions.

Conclusions

Although already prevalent, the number of patients conducting their own web-based medical research increases every year [13,15]; yet, not enough is known about what makes a medical website persuasive to the average patient. Our research identifies three website traits (authorship, format, and tone) relevant to patient credibility assessments and provides evidence about how these traits influence patients. Furthermore, we extend the conversation about medical skepticism and show how patients with high levels of medical skepticism may interpret website traits to assess medical advice credibility differently.

Conflicts of Interest

None declared.

References

1. Funk C, Hefferon M, Kennedy B, Johnson C. Americans often trust practitioners more than researchers but are skeptical about scientific integrity. Pew Research Center. 2019. URL: https://www.pewresearch.org/science/2019/08/02/americans-often-trust-practitioners-more-than-researchers-but-are-skeptical-about-scientific-integrity/ [accessed 2022-01-31]
2. Gallagher J, Lawrence HY. Rhetorical appeals and tactics in New York Times comments about vaccines: qualitative analysis. J Med Internet Res 2020;22(12):e19504. [doi: 10.2196/19504] [Medline: 33275110]
3. LaCour M, Davis T. Vaccine skepticism reflects basic cognitive differences in mortality-related event frequency estimation. Vaccine 2020;38(21):3790-3799. [doi: 10.1016/j.vaccine.2020.02.052] [Medline: 32169393]
4. Waters EA, Ball L, Gehlert S. "I don't believe it." Acceptance and skepticism of genetic health information among African-American and White smokers. Soc Sci Med 2017;184:153-160 [FREE Full text] [doi: 10.1016/j.socscimed.2017.04.053] [Medline: 28527373]
5. Callahan LF, Freburger JK, Mielenz TJ, Wiley-Exley EK. Medical skepticism and the use of complementary and alternative health care providers by patients followed by rheumatologists. J Clin Rheumatol 2008;14(3):143-147. [doi: 10.1097/RHU.0b013e31817734cd] [Medline: 18525432]
6. Fiscella K, Franks P, Clancy CM, Doescher MP, Banthin JS. Does skepticism towards medical care predict mortality? Med Care 1999;37(4):409-414. [doi: 10.1097/00005650-199904000-00010] [Medline: 10213021]
7. Rohrer JE, Borders TF. Healthy skepticism. Prev Med 2004;39(6):1234-1237. [doi: 10.1016/j.ypmed.2004.04.038] [Medline: 15539061]
8. Fiscella K, Franks P, Clancy CM. Skepticism toward medical care and health care utilization. Med Care 1998;36(2):180-189. [doi: 10.1097/00005650-199802000-00007] [Medline: 9475472]
9. Bell RA, Grzywacz JG, Quandt SA, Neiberg R, Lang W, Nguyen H, et al. Medical skepticism and complementary therapy use among older rural African-Americans and Whites. J Health Care Poor Underserved 2013;24(2):777-787 [FREE Full text] [doi: 10.1353/hpu.2013.0052] [Medline: 23728044]
10. Jensen JD, Carcioppolo N, King AJ, Bernat JK, Davis L, Yale R, et al. Including limitations in news coverage of cancer research: effects of news hedging on fatalism, medical skepticism, patient trust, and backlash. J Health Commun 2011;16(5):486-503. [doi: 10.1080/10810730.2010.546491] [Medline: 21347947]
11. Kata A. A postmodern Pandora's box: anti-vaccination misinformation on the Internet. Vaccine 2010;28(7):1709-1716. [doi: 10.1016/j.vaccine.2009.12.022] [Medline: 20045099]
12. Kata A. Anti-vaccine activists, Web 2.0, and the postmodern paradigm--an overview of tactics and tropes used online by the anti-vaccination movement. Vaccine 2012;30(25):3778-3789. [doi: 10.1016/j.vaccine.2011.11.112] [Medline: 22172504]
13. Higgins O, Sixsmith J, Barry MM, Domegan C. A literature review on health information-seeking behaviour on the web: a health consumer and health professional perspective. Insights into health communication. Stockholm: European Centre for Disease Prevention and Control; 2011.
14. Wang L, Wang J, Wang M, Li Y, Liang Y, Xu D. Using Internet search engines to obtain medical information: a comparative study. J Med Internet Res 2012;14(3):e74 [FREE Full text] [doi: 10.2196/jmir.1943] [Medline: 22672889]
15. Fox S, Duggan M. Health online 2013. Pew Research Center. 2013. URL: https://www.pewinternet.org/2013/01/15/health-online-2013/ [accessed 2022-01-31]
16. Podichetty VK, Booher J, Whitfield M, Biscup RS. Assessment of internet use and effects among healthcare professionals: a cross sectional survey. Postgrad Med J 2006;82(966):274-279 [FREE Full text] [doi: 10.1136/pgmj.2005.040675] [Medline: 16597816]
17. Morris CA, Avorn J. Internet marketing of herbal products. JAMA 2003;290(11):1505-1509. [doi: 10.1001/jama.290.11.1505] [Medline: 13129992]
18. Purcell GP, Wilson P, Delamothe T. The quality of health information on the internet. BMJ 2002;324(7337):557-558 [FREE Full text] [doi: 10.1136/bmj.324.7337.557] [Medline: 11884303]
19. McBride JA, Carson CC, Coward RM. Readability, credibility and quality of patient information for hypogonadism and testosterone replacement therapy on the Internet. Int J Impot Res 2017;29(3):110-114. [doi: 10.1038/ijir.2017.4] [Medline: 28228646]
20. Silberg WM, Lundberg GD, Musacchio RA. Assessing, controlling, and assuring the quality of medical information on the Internet: Caveant lector et viewor--let the reader and viewer beware. JAMA 1997;277(15):1244-1245. [Medline: 9103351]
21. Eastin MS. Credibility assessments of online health information: the effects of source expertise and knowledge of content. J Comput Mediat Commun 2001;6(4). [doi: 10.1111/j.1083-6101.2001.tb00126.x]
22. Chen YY, Li CM, Liang JC, Tsai CC. Health information obtained from the Internet and changes in medical decision making: questionnaire development and cross-sectional survey. J Med Internet Res 2018;20(2):e47 [FREE Full text] [doi: 10.2196/jmir.9370] [Medline: 29434017]

23. Amante DJ, Hogan TP, Pagoto SL, English TM, Lapane KL. Access to care and use of the Internet to search for health information: results from the US National Health Interview Survey. J Med Internet Res 2015;17(4):e106. [doi: 10.2196/jmir.4126] [Medline: 25925943]
24. Alduraywish SA, Altamimi LA, Aldhuwayhi RA, AlZamil LR, Alzeghayer LY, Alsaleh FS, et al. Sources of health information and their impacts on medical knowledge perception among the Saudi Arabian population: cross-sectional study. J Med Internet Res 2020;22(3):e14414. [doi: 10.2196/14414] [Medline: 32191208]
25. Palen TE, Ross C, Powers JD, Xu S. Association of online patient access to clinicians and medical records with use of clinical services. JAMA 2012;308(19):2012-2019. [doi: 10.1001/jama.2012.14126] [Medline: 23168824]
26. Ford JD, Ford LW. The role of conversations in producing intentional change in organizations. Acad Manage Rev 1995;20(3):541-570. [doi: 10.2307/258787]
27. Bajracharya AS, Crotty BH, Kowoloff HB, Safran C, Slack WV. Patient experience with family history tool: analysis of patients' experience sharing their family health history through patient-computer dialogue in a patient portal. J Am Med Inform Assoc 2019;26(7):603-609. [doi: 10.1093/jamia/ocz008] [Medline: 30946464]
28. Wang X, Parameswaran S, Bagul DM, Kishore R. Can online social support be detrimental in stigmatized chronic diseases? A quadratic model of the effects of informational and emotional support on self-care behavior of HIV patients. J Am Med Inform Assoc 2018;25(8):931-944 [FREE Full text] [doi: 10.1093/jamia/ocy012] [Medline: 29684190]
29. Laranjo L, Arguel A, Neves AL, Gallagher AM, Kaplan R, Mortimer N, et al. The influence of social networking sites on health behavior change: a systematic review and meta-analysis. J Am Med Inform Assoc 2015;22(1):243-256. [doi: 10.1136/amiajnl-2014-002841] [Medline: 25005606]
30. Yan L, Tan Y. Feeling blue? Go online: an empirical study of social support among patients. Inf Syst Res 2014;25(4):690-709. [doi: 10.1287/isre.2014.0538]
31. Erfani SS, Blount Y, Abedin B. The influence of health-specific social network site use on the psychological well-being of cancer-affected people. J Am Med Inform Assoc 2016;23(3):467-476. [doi: 10.1093/jamia/ocv170] [Medline: 26911816]
32. Dredze M. How social media will change public health. IEEE Intell Syst 2012;27(4):81-84. [doi: 10.1109/mis.2012.76]
33. Eysenbach G. Medicine 2.0: social networking, collaboration, participation, apomediation, and openness. J Med Internet Res 2008;10(3):e22 [FREE Full text] [doi: 10.2196/jmir.1030] [Medline: 18725354]
34. Hesse BW, Hansen D, Finholt T, Munson S, Kellogg W, Thomas JC. Social participation in health 2.0. Computer (Long Beach Calif) 2010;43(11):45-52 [FREE Full text] [doi: 10.1109/MC.2010.326] [Medline: 21379365]
35. Harris PR, Sillence E, Briggs P. Perceived threat and corroboration: key factors that improve a predictive model of trust in internet-based health information and advice. J Med Internet Res 2011;13(3):e51 [FREE Full text] [doi: 10.2196/jmir.1821] [Medline: 21795237]
36. Gefen D, Karahanna E, Straub DW. Trust and TAM in online shopping: an integrated model. MIS Q 2003;27(1):51-90. [doi: 10.2307/30036519]
37. DeTora LM. The spectre of ghostwriting: eroding public trust in physicians, clinical trial integrity and biomedical authorship. Int J Clin Pract 2016;70(7):630-633. [doi: 10.1111/ijcp.12839] [Medline: 27354171]
38. Yang Q, Beatty M. A meta-analytic review of health information credibility: belief in physicians or belief in peers? Health Inf Manag 2016;45(2):80-89. [doi: 10.1177/1833358316639432] [Medline: 27105476]
39. Hu Y, Shyam Sundar S. Effects of online health sources on credibility and behavioral intentions. Commun Res 2009;37(1):105-132. [doi: 10.1177/0093650209351512]
40. Giustini D. How Web 2.0 is changing medicine. BMJ 2006;333(7582):1283-1284 [FREE Full text] [doi: 10.1136/bmj.39062.555405.80] [Medline: 17185707]
41. Murugesan S. Understanding Web 2.0. IT Prof 2007;9(4):34-41. [doi: 10.1109/MITP.2007.78]
42. Cyr D, Head M, Lim E, Stibe A. Using the elaboration likelihood model to examine online persuasion through website design. Inf Manag 2018;55(7):807-821. [doi: 10.1016/j.im.2018.03.009]
43. Barcelos RH, Dantas DC, Sénécal S. Watch your tone: how a brand's tone of voice on social media influences consumer responses. J Interact Mark 2018;41(C):60-80. [doi: 10.1016/j.intmar.2017.10.001]
44. Boas TC, Christenson DP, Glick DM. Recruiting large online samples in the United States and India: Facebook, Mechanical Turk, and Qualtrics. Political Sci Res Methods 2018;8(2):232-250. [doi: 10.1017/psrm.2018.28]
45. Akdeniz B, Calantone RJ, Voorhees CM. Effectiveness of marketing cues on consumer perceptions of quality: the moderating roles of brand reputation and third-party information. Psychol Mark 2013;30(1):76-89. [doi: 10.1002/mar.20590]
46. Lagasse J. Why branding in healthcare is more important than ever, and how to connect with consumers. Healthcare Finance News. 2018. URL: https://www.healthcarefinancenews.com/news/why-branding-healthcare-more-important-ever-and-how-connect-consumers [accessed 2020-07-28]
47. Ferreira G, Traeger AC, Machado G, O'Keeffe M, Maher CG. Credibility, accuracy, and comprehensiveness of internet-based information about low back pain: a systematic review. J Med Internet Res 2019;21(5):e13357 [FREE Full text] [doi: 10.2196/13357] [Medline: 31066689]
48. Turner AM, Osterhage KP, Taylor JO, Hartzler AL, Demiris G. A closer look at health information seeking by older adults and involved family and friends: design considerations for health information technologies. AMIA Annu Symp Proc 2018;2018:1036-1045 [FREE Full text] [Medline: 30815147]

49. Phua J, Ahn SJ. Explicating the 'like' on Facebook brand pages: the effect of intensity of Facebook use, number of overall 'likes', and number of friends' 'likes' on consumers' brand outcomes. J Mark Commun 2014;22(5):544-559. [doi: 10.1080/13527266.2014.941000]
50. Estill J. Shorter is better. Copyblogger. 2008. URL: https://copyblogger.com/shorter-is-better/ [accessed 2020-07-28]
51. Yu CH. Web page enhancement on desktop and mobile browsers. Massachusetts Institute of Technology Libraries. 2012. URL: https://dspace.mit.edu/handle/1721.1/79216 [accessed 2020-07-28]
52. Ibrahim N, Shiratuddin MF, Wong KW. Persuasion techniques for tourism website design. In: Proceedings of the international conference on e-technologies and business on the Web. 2013 Presented at: EBW '13; May 7-9, 2013; Bangkok URL: https://researchrepository.murdoch.edu.au/id/eprint/22861/
53. Anderson R, Lion J, Anderson O. Two decades of health services research: social survey trends in use and expenditures. Cambridge: Ballinger Publishing Company; 1976.
54. Gilmour J, Hanna S, Chan H, Strong A, Huntington A. Engaging with patient online health information use: a survey of primary health care nurses. SAGE Open 2014;4(3):215824401455061. [doi: 10.1177/2158244014550617]
55. Tan SS, Goonawardene N. Internet health information seeking and the patient-physician relationship: a systematic review. J Med Internet Res 2017;19(1):e9 [FREE Full text] [doi: 10.2196/jmir.5729] [Medline: 28104579]
56. Folger JK, Nam CB, U.S. Department of Commerce, Census Bureau, U.S. Census of Population. Rates of high school completion and bachelor's degree attainment among persons age 25 and over, by race/ethnicity and sex: selected years, 1910 through 2018. National Center for Education Statistics. 2019. URL: https://nces.ed.gov/programs/digest/d18/tables/dt18_104.10.asp [accessed 2022-01-31]
57. DeVoe JE, Wallace LS, Fryer Jr GE. Patient age influences perceptions about health care communication. Fam Med 2009;41(2):126-133 [FREE Full text] [Medline: 19184691]
58. Nabalamba A, Millar WJ. Going to the doctor. Health Rep 2007;18(1):23-35 [FREE Full text] [Medline: 17441441]
59. Borders TF, Rohrer JE, Xu KT, Smith DR. Older persons' evaluations of health care: the effects of medical skepticism and worry about health. Health Serv Res 2004;39(1):35-52. [doi: 10.1111/j.1475-6773.2004.00214.x] [Medline: 14965076]
60. McGrath JE. Dilemmatics: the study of research choices and dilemmas. Am Behav Sci 1981;25(2):179-210. [doi: 10.1177/000276428102500205]
61. Bach RL, Wenz A. Studying health-related internet and mobile device use using web logs and smartphone records. PLoS One 2020;15(6):e0234663. [doi: 10.1371/journal.pone.0234663] [Medline: 32530937]
62. Willis E. Patients' self-efficacy within online health communities: facilitating chronic disease self-management behaviors through peer education. Health Commun 2016;31(3):299-307. [doi: 10.1080/10410236.2014.950019] [Medline: 26325224]
63. Willis E, Royne MB. Online health communities and chronic disease self-management. Health Commun 2017;32(3):269-278. [doi: 10.1080/10410236.2016.1138278] [Medline: 27218836]
64. Stellefson M, Chaney B, Barry AE, Chavarria E, Tennant B, Walsh-Childers K, et al. Web 2.0 chronic disease self-management for older adults: a systematic review. J Med Internet Res 2013;15(2):e35 [FREE Full text] [doi: 10.2196/jmir.2439] [Medline: 23410671]

Abbreviations

OLS: ordinary least squares

Edited by A Mavragani; submitted 31.03.21; peer-reviewed by H Gewald, S Pandey, YY Chen; comments to author 10.05.21; revised version received 23.05.21; accepted 12.12.21; published 18.02.22.

Please cite as:
Claggett J, Kitchens B, Paino M, Beisecker Levin K
The Effects of Website Traits and Medical Skepticism on Patients' Willingness to Follow Web-Based Medical Advice: Web-Based Experiment
J Med Internet Res 2022;24(2):e29275
URL: https://www.jmir.org/2022/2/e29275
doi:10.2196/29275
PMID:35179506

Original Paper

Real-world Implementation of a Smartphone-Based Psychoeducation Program for Bipolar Disorder: Observational Ecological Study

Aitana García-Estela[1,2], MSc; Jordi Cantillo, BSc, MSc; Natalia Angarita-Osorio[1,2], MSc; Estanislao Mur-Milà[1,2,3], MD; Gerard Anmella[4,5], MD; Víctor Pérez[1,2,3,5], MD, PhD; Eduard Vieta[4,5], MD, PhD; Diego Hidalgo-Mazzei[4,5,6], MD, PhD; Francesc Colom[1,3,5,7], MSc, PhD

[1]Mental Health Research Group, Hospital del Mar Medical Research Institute (IMIM), Barcelona, Spain

[2]Department of Psychiatry and Forensic Medicine, School of Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain

[3]Institute of Neuropsychiatry and Addictions, Hospital del Mar, Parc de Salut Mar, Barcelona, Spain

[4]Bipolar and Depressive Disorders Unit, Department of Psychiatry and Psychology, Institute of Neuroscience, Hospital Clínic, Universitat de Barcelona, Barcelona, Spain

[5]Centre for Biomedical Research in Mental Health Network (CIBERSAM), Madrid, Spain

[6]Centre for Affective Disorders, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom

[7]Department of Basic, Evolutive and Education Psychology, Faculty of Psychology, Universitat Autònoma de Barcelona, Barcelona, Spain

Corresponding Author:
Francesc Colom, MSc, PhD
Mental Health Research Group
Hospital del Mar Medical Research Institute (IMIM)
Office 202, PRBB Building
Doctor Aiguader, 88
Barcelona, 08003
Spain
Phone: 34 933160400 ext 1493
Email: fcolom@imim.es

Abstract

Background: SIMPLe is an internet - delivered self - management mobile app for bipolar disorder (BD) designed to combine technology with evidence-based interventions and facilitate access to psychoeducational content. The SIMPLe app was launched to the real world to make it available worldwide within the context of BD treatment.

Objective: The main aims of this study are as follows: to describe app use, engagement, and retention rates based on server data; to identify patterns of user retention over the first 6-month follow-up of use; and to explore potential factors contributing to discontinuation of app use.

Methods: This was an observational ecological study in which we pooled available data from a real-world implementation of the SIMPLe app. Participation was open on the project website, and the data-collection sources were a web-based questionnaire on clinical data and treatment history administered at inclusion and at 6 months, subjective data gathered through continuous app use, and the use patterns captured by the app server. Characteristics and engagement of regular users, occasional users, and no users were compared using 2-tailed *t* tests or analysis of variance or their nonparametric equivalent. Survival analysis and risk functions were applied to regular users' data to examine and compare use and user retention. In addition, a user evaluation analysis was performed based on satisfaction, perceived usefulness, and reasons to discontinue app use.

Results: We included 503 participants with data collected between 2016 and 2018, of whom 77.5% (n=390) used the app. Among the app users, 44.4% (173/390) completed the follow-up assessment, and data from these participants were used in our analyses. Engagement declined gradually over the first 6 months of use. The probability of retention of the regular users after 1 month of app use was 67.4% (263/390; 95% CI 62.7%-72.4%). Age (*P*=.002), time passed since illness onset (*P*<.001), and years since diagnosis of BD (*P*=.048) correlate with retention duration. In addition, participants who had been diagnosed with BD for longer used the app on more days (mean 97.73, SD 69.15 days; *P*=.002) than those who had had a more recent onset (mean 66.49, SD 66.18 days; *P*=.002) or those who had been diagnosed more recently (mean 73.45, SD 66 days; *P*=.01).

Conclusions: The user retention rate of the app decreased rapidly after each month until reaching only one-third of the users at 6 months. There exists a strong association between age and app engagement of individuals with BD. Other variables such as years lived with BD, diagnosis of an anxiety disorder, and taking antipsychotics seem relevant as well. Understanding these associations can help in the definition of the most suitable user profiles for predicting trends of engagement, optimization of app prescription, and management.

(J Med Internet Res 2022;24(2):e31565) doi:10.2196/31565

KEYWORDS

bipolar disorder; psychoeducation; smartphone; app; SIMPLe; Intervention; mobile phone

Introduction

Background

Globally, an estimated 46 million people have been diagnosed with bipolar disorder (BD) [1]. Besides behavioral changes occurring during mood episodes, BD has a serious impact on psychosocial functioning, cognition, quality of life, and survival rate of individuals with this condition [2,3]. Although ranges vary dramatically because of methodological differences among studies, there is agreement that people with BD are between 9 and 30 times more likely to die from suicide than someone without this condition [4,5].

Although the fundamental treatment for BD relies basically on psychopharmacology, some adjuvant psychological interventions have been shown to improve the long-term outcomes of this disease [6]. Among psychological interventions, psychoeducation programs have proven to be a cost-effective approach to prevent episodes by helping patients to improve adherence, embrace healthy habits, and learn to recognize the prodromes and symptoms of upcoming episodes. It is obvious that efficacious treatments only work for patients who can receive them [7]; yet, these psychological interventions are scarce and difficult to access for most patients [8].

e-Mental health—the delivery of mental health–related tools through the internet and related technologies [9]—is one of the most promising strategies to address this access gap, relieving overburdened mental health services and increasing the services' cost-efficiency while maintaining its quality. Although internet-based interventions for mental health are still relatively new, the number of solutions such as web-based platforms, smartphone apps, and wearables is rapidly growing globally [10], with their actual acceptability by patients tending to be increasingly higher [11].

However, the strong desire to proliferate e-mental health solutions has not been translated into a transformation in the delivery of mental health care because there is little available evidence of uptake of mental health apps [12]. Moreover, despite the increasing number of e-mental health apps available in app stores, few are properly validated in a scientific process; this lack of validation could jeopardize the safety and health of potential users [13,14]. In addition, the quality of the content of the apps for BD in the app stores does not live up to expectations because most of them do not rely on the available best practice clinical guidelines [15]. Thus, it remains a challenge to implement platforms developed within evidence-based–practice frames that, simultaneously, have been subjected to efficacy tests [16,17].

Within this framework, the SIMPLe project was designed to leverage the potential of combining technology with evidence-based interventions by developing and evaluating an internet - delivered self - management mobile app (SIMPLe 1.5) for BD in addition to standard treatment. The SIMPLe 1.5 app collects information about potential symptoms, with the advantage of providing users with personalized psychoeducational messages and alerts that are tailored to specific needs. The app is based on group psychoeducation, a well-established and evidence-based care-focused psychological intervention that addresses relevant issues of self-management for BD, such as identification and management of early warning signs, lifestyle, and treatment adherence [18,19].

Up until now, the SIMPLe app has proved acceptable to users and has shown interesting and optimistic results: a high retention rate was attained in a 3-month feasibility study and positive outcomes regarding satisfaction were found in a naturalistic implementation feasibility study [20-22]. In addition, some potential improvements in terms of biological rhythms and medication adherence were suggested by post hoc analyses [23]. It is worth mentioning that pharmacological treatment adherence is a particularly complex issue in BD because more than 50% of the patients are estimated to be nonadherent fully to the prescribed doses of medication [24].

Considering that the ultimate aim of the SIMPLe project is to extend and facilitate access to psychoeducational content through the SIMPLe app to all potential users, wide and free access to the SIMPLe 1.5 app around the Spanish-speaking world was offered. This way, we had the opportunity to routinely collect implementation data on use in a real-world setting and naturalistic condition.

A previous report on the OpenSIMPLe study presented partial data demonstrating the high dropout rates when a psychoeducation smartphone-based intervention for individuals with BD is offered openly [22]. We hereby present the results of the whole sample of SIMPLe users throughout the study. This research differs from our previous partial sample in terms of both sample size and follow-up time frame. Statistical analyses of the previous partial sample only included descriptive, pre–post, and logistic regression tests without considering the temporal retention period as we did in this paper.

In this paper, we focus on exploratory analyses that aim to investigate in depth the relationships among variables that may predict overall engagement as well as retention rates mainly by

means of survival analyses. More specifically, we intend to shed some light on the ways in which the SIMPLe app engagement and user retention patterns are influenced by individual variables, including sociodemographic and clinical data.

This study is based on an ecological experimental implementation of an e-mental health resource. Thus, the reader may miss the usual randomized controlled studies' constraints such as the lack of sample size calculation (which, by definition, would be *the bigger the better*) or some control measures. However, this study reflects the real day-to-day problems faced by a mental health app when launched to the app stores to be used by the target population.

Objectives

The main aims of this study are to (1) describe app use, engagement, and retention rates based on server data; (2) identify patterns of user retention over the first 6-month follow-up of use; and (3) explore potential factors contributing to the risk of discontinuing app use.

The expectations are that these exploratory analyses will help to confirm preliminary use data of the SIMPLe app and understand user retention rates as well as the ways in which users self-manage BD in a real-world setting. In addition, we hope that the results will provide our research colleagues with relevant insights into the interplay, dynamics, and predictive factors of user engagement with mental health apps at the time of their implementation in real-world conditions.

Methods

Procedure and Participants

Participation availability was open on the project website [25] to anyone with BD who was fluent in Spanish. Specific app characteristics and functionality offered for the SIMPLe app have been published elsewhere [20,21]. In brief, the app consisted of a daily graphic 5-item screening test (mood, energy, sleep time, medication adherence, and irritability) and a weekly, more comprehensive Yes or No test, considering all Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, criteria for manic and depressive episodes. The resulting mood assessments were displayed in a chart on the home screen. On the basis of the information collected, a daily pop-up notification prompted the user to read a short psychoeducational message providing brief information or advice about how to deal with specific situations to avoid relapses. Each message was extracted from a library of more than 500 messages categorized according to different clinical situations based on a published psychoeducation manual and lay-language books on BD written by 2 of the authors (FC and EV). Additional optional modules available in the app were personalized medication reminders, prodromal symptoms, and mood-chart sharing.

Following a real-world naturalistic approach, no active recruitment strategy or advertisement was used. Potential participants approached the study voluntarily through the website [25] (more detailed information can be found elsewhere [22]). Potential participants needed to meet the following inclusion criteria: (1) aged ≥18 years; (2) having a psychiatrist-confirmed diagnosis of BD before entry; (3) pharmacological treatment for BD provided by a psychiatrist; (4) owning and using daily an Android or iOS smartphone; (5) fluency in Spanish; (6) an active email account; and (7) standard cutoff score of 7, co-occurrence of at least two symptoms, and moderate or severe impairment on a modified version of the Mood Disorder Questionnaire (MDQ) [26,27]. The participants' answers were reviewed by a psychiatrist to assess consistency, after which they were informed whether they were eligible for the study.

To prevent duplicate use and potential misuse, the possibility of completing the questionnaire multiple times from the same IP address was blocked. Web-based technical support was provided to the app users, if needed, through email.

Data from the sample were drawn from 503 SIMPLe users who had provided informed consent and completed the app's onboarding questionnaire between May 2016 and June 2018. Of the 503 participants, 390 (77.5%) used the app, and data from these participants were used in the analyses.

All procedures contributing to this work complied with the ethical standards of national and international guidelines and the basic principles of protection of dignity and human rights, as stated in the Declaration of Helsinki (64th General Assembly, Fortaleza, Brazil, October 2013), and were conducted according to current regulations. The ethical committees from both Hospital Clínic of Barcelona (HCB/2016/0403) and the Hospital del Mar Medical Research Institute (2016/6764/I) approved the protocol.

Assessments

In all, two sources of data were used: (1) a web-based form administered at program inclusion and at 6 months and (2) subjective data gathered through continuous app use and the use patterns captured by the app server.

Psychometric External Assessments

Sociodemographic data as well as illness and treatment history were collected at program inclusion through a web-based form from participants who had provided informed consent. The number of hospitalizations and suicide attempts as well as treatment history during the past 6 months were also collected 6 months after inclusion.

The Spanish validated version of the 5-item World Health Organization Well-being Index [28] was used to assess mental well-being at baseline and 6 months later at the final follow-up.

App-Derived Assessments

Subjective Information

The subjective information assessed was as follows:

1. Self-reported mood, sleep, medication adherence, and energy: The app prompts users to answer 5 daily slider screening tests on mood, energy, sleep time, irritability, and drug-treatment adherence. The daily scores appearing in the chart are the results of an algorithm, which was previously tested during the development phase [20]. In addition, a more comprehensive test, considering Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition,

criteria for manic and depressive episodes, including suicide thoughts, was prompted weekly.

2. Self-reported usability: The app was evaluated by users who made it through to complete follow-up assessment. The users' perception of the usability of the app was measured through the System Usability Scale (SUS) [29], a 10-item questionnaire with 5 response options (from strongly agree to strongly disagree) that allows evaluation of products and services. Interpretation of the raw scores was achieved by converting them into percentile ranks [30] and associating them with adjectives [31].
3. Satisfaction and perceived helpfulness: Satisfaction and perceived helpfulness of each subcomponent of the app were assessed in the follow-up questionnaire through Likert scales after 6 months of program initiation. Suggestions and comments regarding the app were also registered.

Mobile Terminal–Mined Information

User retention, app use, and engagement data were constantly recorded at the servers over the study duration, reflecting continuously and in detail app use and engagement. User retention was defined as the proportion of participants who used the app for the entire duration of the study and completed the 6-month follow-up assessment. The SIMPLe app has multiple components, three of which we consider the core active ingredients: the daily and weekly tests and the psychoeducational messages. To determine retention, we considered the user to be active if we registered data in the servers from these 3 interactions, and we considered the user to have discontinued participation if there was a lack of data from these variables in the server for >1 month.

Engagement with mobile apps is considered a multidimensional construct, and different definitions can be used or combined to measure it. In this study, engagement was understood as the ability of an app to engage users and sustain user interactions and it was assessed through indicators such as usability, acceptability, and feasibility [32]. In this case, engagement was calculated based on the weekly percentage of completed tasks (ie, answering daily and weekly tests and reading the daily psychoeducational messages).

Design

This is an observational ecological study in which we pooled available data from a real-world implementation of the SIMPLe app.

The rationale for the OpenSIMPLe study and detailed methods have been published elsewhere [22]; the methods are briefly outlined here.

Statistical Analysis

Smartphone app data (ie, participants' mood ratings, manic or hypomanic and depressive symptoms, and details of their use of the app) were downloaded directly from the servers. Likewise, the users' baseline and follow-up responses at both baseline and follow-up web-based questionnaires were retrieved from the servers. All analyses were run using SPSS software (version 26.0; IBM Corp) and R statistical package for Windows (version 4.0.2; The R Foundation for Statistical Computing).

Initially, basic descriptive statistics of sociodemographic and clinical variables were run, including age, sex, marital status, family psychiatric history, follow-up time, number of episodes, substance abuse, and comorbid medical and psychiatric diagnoses. Continuous variables have been described based on the mean and SD; the median and the 25th and 75th percentiles have also been used in comparative analyses of the time spent using the SIMPLe 1.5 app. We defined categorical variables in terms of the number and percentage of users per response category. In addition, statistical techniques were used to confirm assumptions of the statistical tests before carrying out parametric tests to compare means and proportions. When the assumptions of parametric tests were violated, nonparametric tests were used.

We performed a comparative analysis of the variables among groups using a 2-tailed *t* test or analysis of variance on continuous variables or their nonparametric equivalent; the Wilcoxon test or the Kruskal–Wallis test was performed depending on the inherent characteristics of the variables analyzed; the Mann–Whitney U test was carried out when dependent variables were ordinal; and the chi-square test was performed when analyzing categorical variables.

The Spearman correlation coefficient (ρ) was used to examine correlations between SIMPLe 1.5 app time use and the different continuous variables.

A subsample of users with *engagement* ≥12% was selected for the analysis of the use time of the SIMPLe 1.5 app to determine a minimum frequency of use and avoid overestimating the use time. A homogeneity analysis was performed between the selected subsample and the rest of the participants using the chi-square test for categorical variables and the Kolmogorov–Smirnov Z test for continuous variables.

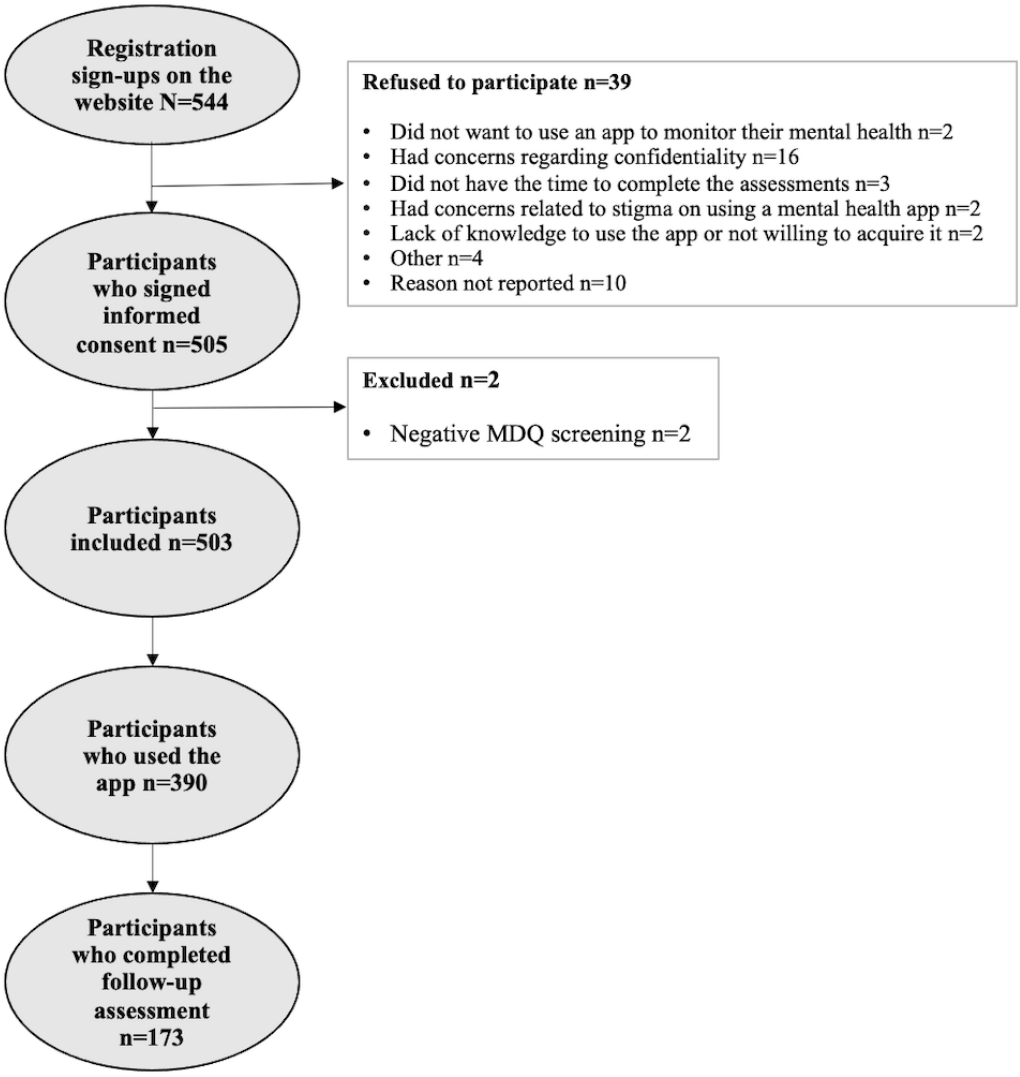
Survival analysis was used on the selected subsample to examine the time spent using the SIMPLe 1.5 app because the length of the follow-up period was variable and there were participants who did not experience the event *quit using the app* during the 6-month follow-up. Estimates of survival and risk functions of the time use of the app were calculated by applying the Kaplan–Meier method. We used the log-rank test to compare various survival distributions and the Cox proportional-hazards model validated by Schoenfeld residual analysis to assess risks related to the survival of the app users. A sensitivity analysis was performed by repeating the survival analysis with all users to evaluate the effect of the selection of the subsample in the study of the use time. Results were considered significant with 2-sided *P* values <.05.

Results

Descriptive Analysis

Figure 1 depicts the number of participants registered on the OpenSIMPLe website who were initially interested in using the app alongside those included in the statistical analysis. The reasons reported by users who did not want to participate in the study are also listed. Finally, potential participants who were excluded are described, as well as the number of participants who actually used the app and those who responded to the 6-month follow-up assessment.

Figure 1. Flowchart of participants included in the statistical analysis. MDQ: Mood Disorder Questionnaire.



Sociodemographic and Clinical Characteristics

The mean age of users was 34.74 (SD 10.48) years, and most (264/390, 67.7%) of them were women. The most frequent ethnicity was Latin American (266/390, 68.2%), with high education levels (241/390, 61.8%). A high percentage of the sample was employed at the time of study entry (156/390, 40%), whereas only 17.7% (69/390) were either on temporary or permanent disability leave. Regarding housing conditions, 33.8% (132/390) of the participants lived in the parental household and more than half of the sample reported living independently, either owning (117/390, 30%) or renting their current house or flat (90/390, 23.1%). Sociodemographic variables of the SIMPLe app users are described in Table 1.

Regarding the clinical variables, a mean disorder duration of 13.23 (SD 9.97) years was identified; 49.7% (194/390) of the users stated that they had experienced ≥10 depressive episodes; and 33.3% (130/390) reported ≥10 manic or hypomanic episodes. Most of the participants who used the app were receiving treatment with at least one mood stabilizer (353/390, 90.5%) and at least one antipsychotic (252/390, 64.6%), whereas almost half (193/390, 49.5%) of the participants were receiving at least one antidepressant. Furthermore, 71.5% (279/390) of the participants were receiving some kind of psychological treatment. The clinical variables collected at baseline are described in Table 2.

Table 1. Baseline sociodemographic characteristics of participants (N=390).

Characteristic	Value
Gender, female, n (%)	264 (67.7)
Ethnicity, n (%)	
African	2 (0.5)
White	119 (30.5)
Latin American	266 (68.2)
Asian	2 (0.5)
Other	1 (0.3)
Age (years), mean (SD)	34.74 (10.48)
Marital status, n (%)	
Single	192 (49.2)
Married	81 (20.8)
Cohabitation	50 (12.8)
Divorced or separated	54 (13.8)
Widowed	1 (0.3)
Other	12 (3.1)
Housing status, n (%)	
Shared home	43 (11)
Tenant	90 (23.1)
Homeowner	117 (30)
Parental home	132 (33.8)
Residence or institution	8 (2.1)
Completed studies, n (%)	
None	1 (0.3)
Primary education	8 (2.1)
Secondary education	67 (17.2)
A-level or general certificate of education	73 (18.7)
Vocational education and training or certificate of higher education or higher national diploma	95 (24.4)
Bachelor's degree	101 (25.9)
Graduate certificate or postgraduate diploma or master's degree	45 (11.5)
Employment status, n (%)	
Unemployed	78 (20)
Student	81 (20.8)
Employed	156 (40)
Retired	10 (2.6)
Temporary disability leave	35 (9)
Permanent disability leave	30 (7.7)
Country, n (%)	
Spain	130 (33.3)
Chile	76 (19.5)
Argentina	66 (16.9)
Mexico	25 (6.4)

Characteristic	Value
Colombia	23 (5.9)
Guatemala	12 (3.1)
Brazil	9 (2.3)
Other	49 (12.6)

Table 2. Baseline clinical variables of app users (N=390).

Illness course	Value
Years since onset, mean (SD)	13.23 (9.97)
Years since diagnosis of bipolar disorder, mean (SD)	6.4 (6.55)
Depressive episodes, n (%)	
0-4	110 (28.2)
5-9	86 (22.1)
≥10	194 (49.7)
Manic or hypomanic episodes, n (%)	
0-4	143 (36.7)
5-9	117 (30)
≥10	130 (33.3)
Previous hospital admissions because of an episode, n (%)	
None	185 (47.4)
1-2	135 (34.6)
≥2	70 (17.9)
Suicide attempts, n (%)	
None	156 (40)
1-2	142 (36.4)
≥2	92 (23.6)
Treatment setting, n (%)	
Public health network	145 (37.2)
Private health network	184 (47.2)
Both	61 (15.6)
Past psychological treatment, n (%)	
None	39 (10)
Yes, individual psychotherapy	260 (66.7)
Yes, group psychotherapy	9 (2.3)
Yes, individual and group psychotherapy	82 (21)
Current psychological treatment, n (%)	
None	111 (28.5)
Individual psychotherapy	230 (59)
Group psychotherapy	15 (3.8)
Individual and group psychotherapy	34 (8.7)
Current pharmacological treatment, n (%)	
Mood stabilizer	353 (90.5)
Antipsychotic	252 (64.6)
Antidepressant	193 (49.5)
Anxiolytic	183 (46.9)

Engagement

The 503 participants included in the study can be divided into three broad categories based on their app use: no users (never used the app), occasional users (engagement <12%), and regular users (engagement ≥12%). Of the 503 participants, 113 (22.5%) were no users, 357 (70.9%) were regular users, and 33 (6.6%) were occasional users. In addition, among the participants who used the app, 44.4% (173/390) completed the follow-up assessment too.

We analyzed the number of days containing any kind of record in the app from users over the 6-month follow-up period. Monthly progress of regular users' engagement declined gradually over the first 6 months. The highest engagement was observed in the first month (mean 0.74, SD 0.20); in the second month, it dropped sharply. At 6 months, the users had a mean engagement of 0.39 (SD 0.34).

Occasional users used the app a mean of 139.06 (SD 56.80) days, with a use frequency of 2.05 (SD 0.87) days per month, whereas regular users used the app 83.98 (SD 69.95) days, with a use frequency of 19.42 (SD 7.76) days per month. The group of occasional users rarely used the app over long periods of time, which could overestimate use time rates, for example, a participant who used the app at the 3-month follow-up for the last time but only used the app 4 times overall. The estimate that this participant used the app for 3 months may lead us to a use bias. As the SIMPLe app was developed for daily use, we could consider that the interaction of this kind of user with the app is low enough to overestimate the time of use in the statistical analysis. To avoid this bias, we used the variable overall engagement—the percentage of tasks completed compared with those expected to be completed during the time in which they used the app—to identify users who may make us enhance the time-of-use overestimation. Occasional users were ruled out in the survival analysis.

Sensitivity and homogeneity analyses show homogeneous baseline variables across subsets of participants defined by retention duration, with the exception of the *cocaine use* variable, which was significantly greater among the group of no users than among the occasional and regular users. In addition, we performed survival analysis again with all users to assess the effect of the selection of regular users in the estimations of retention duration. The results showed a 7-day increase in the median survival time when we included occasional users in the analysis, whereas the probability of survival only increased by a score of 0.027 (SD 0.007) on average. Therefore, we are confident that selection bias did not occur and that the subsets were representative of the data of the sample.

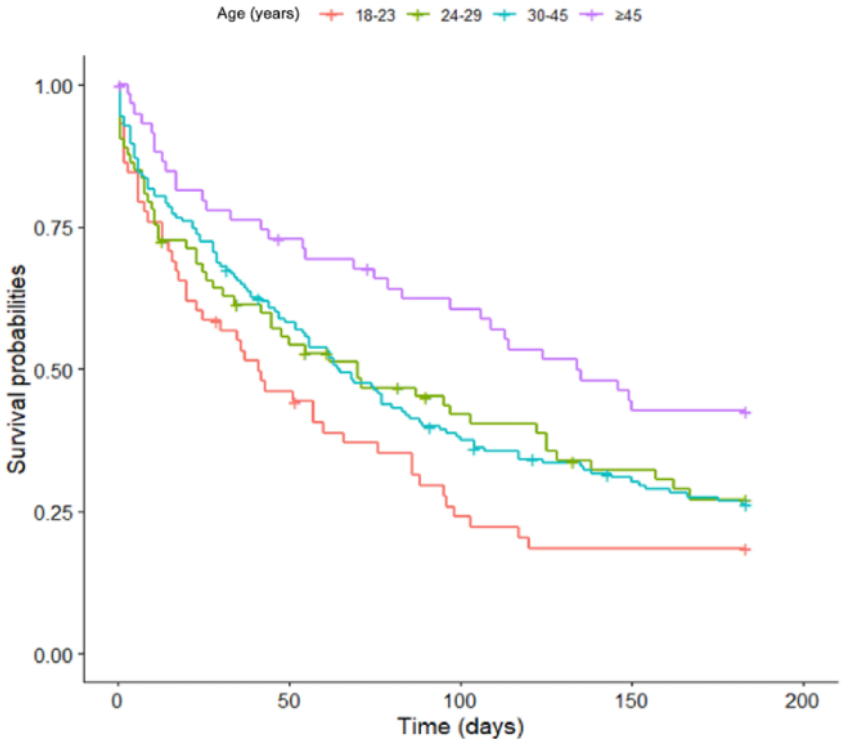
Use and Retention Duration

We analyzed the number of days with records in the app from users over the 6-month follow-up period. It turned out that only 13.8% (54/390) of the users used the SIMPLe app for >100 days. The mean survival time of regular app users was 87.95 (SD 72.08; 95% CI 80.48-95.43) days.

The probability of survival for the 357 participants under consideration for these analyses after 1 month of app use was 67.4% (95% CI 62.7%-72.4%); at 3 months, the probability of survival was 43% (95% CI 38.1%-48.5%); and at the 6-month follow-up assessment, the probability of survival was 28% (95% CI 23.6%-33.2%). The risk of discontinuing app use increased as the days passed: at 3 months, the cumulative risk of discontinuing app use was 83.7%; however, at 6 months, this rose to a cumulative risk of 126.3%.

The correlations between retention duration and the sociodemographic and clinical variables of users at baseline were analyzed. A direct correlation between age and engagement was observed (ρ=0.168; *P*=.002); older users (aged ≥46 years) had higher total app use (mean 109.78, SD 71.42; *P*=.005) than younger users (aged 18-23 years; mean 63.12, SD 63.94; *P*=.005; Figure 2).

Figure 2. Plot of Kaplan–Meier age estimates of survival of participants using the SIMPLe app. The horizontal axis represents the survival time (in days) with records in the app (6 months maximum).



Time passed since illness onset was strongly associated with time of use (ρ=0.194; *P*<.001) and years since diagnosis of BD (ρ=0.106; *P*=.048). In addition, participants who had been diagnosed with BD for longer (>17 years) used the app for more days (mean 97.73, SD 69.15; *P*=.002) than those who had had a more recent onset (mean 66.49, SD 66.18; *P*=.002) or those who had been diagnosed <6.5 years ago (mean 73.45, SD 66; *P*=.01).

We performed log-rank (Mantel–Cox) test analysis to compare survival curves and detect potential factors contributing to the risk of discontinuing app use (Table 3). Variables with a significant contribution were years since onset (χ^2_2=11.7; *P*=.003), years with BD diagnosis (χ^2_1=8.9; *P*=.003), opiate use (χ^2_1=7.9; *P*=.005), age (χ^2_3=12.3; *P*=.006), housing status (χ^2_4=12.3; *P*=.006), employment status (χ^2_5=13.5; *P*=.02), and antipsychotic use (χ^2_1=4.9; *P*=.03).

Survival-curve plots for the variables of interest were produced over the 6-month-long follow-up period and are described in the following paragraphs (Table 4).

Table 3. Log-rank test for overall survival[a].

Variables	Log rank (Mantel–Cox)	
	Chi-square (*df*)	*P* value
Clinical variables		
Comorbid psychiatric disorder	2.5 (1)	.11
Anxiety disorder	3.9 (1)	.04
Personality disorder	2.4 (1)	.12
Substance abuse disorder	0.5 (1)	.46
Eating disorder	2.0 (1)	.15
PTSD[b]	1.1 (1)	.30
Other comorbid psychiatric disorders	2.2 (1)	.13
WHO-5[c]	2.9 (2)	.23
Illness course		
Years since onset of first episode	11.7 (2)	.003
Years diagnosed with bipolar disorder	8.9 (1)	.003
Depressive episodes	4.2 (2)	.12
Manic or hypomanic episodes	4.5 (2)	.10
Hospitalizations because of an episode	2.1 (2)	.35
Suicide attempts	1.9 (2)	.36
Treatment		
Psychotherapy	1.1 (3)	.77
Mood stabilizer	0.2 (1)	.65
Antipsychotic	4.9 (1)	.02
Antidepressant	0.0 (1)	.84
Anxiolytic	0.1 (1)	.73
Electroconvulsive therapy	1.8 (1)	.17

[a]Users with engagement ≥12%.

[b]PTSD: posttraumatic stress disorder.

[c]WHO-5: 5-item World Health Organization Well-being Index.

Table 4. Comparing survival curves.

Characteristic	Regular users[a]	
	Mean, estimate (SE; 95% CI)	Median, estimate (SE; 95% CI)
Age (years)		
18-23	65.72 (8.60; 48.85-82.58)	41.00 (9.90; 21.59-60.40)
24-29	87.11 (8.69; 70.06-104.16)	70.00 (23.12; 24.67-115.32)
30-45	86.26 (5.54; 75.40-97.12)	65.00 (7.28; 50.72-79.27)
≥46	114.83 (9.17; 96.85-132.80)	134.00 (22.24; 90.39-177.60)
Housing status		
Shared home	62.65 (9.91; 43.21-82.08)	50.00 (8.51; 33.31-66.68)
Tenant	98.03 (7.71; 82.91-113.15)	83.00 (16.02; 51.58-114.41)
Homeowner	99.95 (7.34; 85.56-114.34)	109.00 (26.39; 57.26-160.73)
Parental house	77.52 (6.35; 65.05-89.98)	56.00 (10.86; 34.70-77.29)
Residence or institution	96.62 (20.35; 56.72-136.52)	62.00 (60.81; 0.00-181.19)
Employment status		
Unemployed	72.87 (7.74; 57.69-88.05)	55.00 (12.96; 29.58-80.41)
Student	75.02 (7.85; 59.62-90.42)	57.00 (14.14; 29.28-84.71)
Employed	97.40 (6.09; 85.45-109.36)	80.00 (12.27; 55.9-104.06)
Retired	109.40 (24.72; 60.93-157.86)	84.00[b]
Temporary disability leave	101.96 (13.21; 76.06-127.86)	104.00 (61.23; 0.00-224.01)
Permanent disability leave	85.14 (14.47; 56.77-113.52)	69.00 (24.71; 20.55-117.44)
Diagnostic time (years)		
0-6.5	81.14 (4.79; 71.74-90.54)	61.00 (7.45; 46.38-75.61)
>6.5	98.15 (6.36; 85.67-110.63)	94.00 (15.98; 62.67-125.32)
Anxiety disorder		
No	95.07 (4.91; 85.43-104.71)	83.00 (9.17; 65.01-100.98)
Yes	78.19 (5.93; 66.56-89.83)	46.00 (7.96; 30.39-61.61)

[a]Users with engagement ≥12%.

[b]SE and 95% CI are not available.

The app use survivorship of the oldest participants (aged ≥46 years) seems greater than that of the youngest group of users (aged 18-23 years) because the estimated mean was 114.83 (95% CI 96.85-132.80) days for users in the former age range, whereas the use mean was 65.72 (95% CI 48.85-82.58) days for users in the latter group. At 60 days, the probability of survival of the youngest users was 38.9% (95% CI 28%-53.9%); this likelihood increased to 69.4% for the oldest group of participants, with a cumulative risk of 0.923 and 0.36, respectively.

Regarding housing status, we observed that the mean estimation of app use survival of people sharing a house or flat was 62.65 (95% CI 43.21-82.08) days, that is, between 15 and 37 days fewer than users with other housing statuses, suggesting a survival disadvantage. At 60 days, the cumulative incidence risk estimates among users who shared a house were 0.880, whereas they were 0.435 for individuals who lived in residences, 0.488 for tenants, 0.509 for homeowners, and 0.734 for those who lived in the parental home.

Being unemployed seemed to worsen app use survivorship pretty much at all time points because the survival likelihood mean estimation of unemployed participants was 72.87 (95% CI 57.69-88.05) days, that is, lower than any other employment status. At 60 days, the highest survival cumulative risk was that of unemployed participants (0.737), followed by students (0.734), individuals on permanent (0.596) and temporary disability leave (0.573), employers (0.523), and retired people (0.336).

In addition, app use declined faster among participants who had been recently diagnosed (<6.5 years) compared with users who had been diagnosed for a longer period of time; at 60 days, the cumulative risk of app use discontinuation among people who had a recent diagnosis was 0.681, whereas this risk was lower for those who had been diagnosed earlier (0.518). The mean app use estimation of individuals with a more recent diagnosis of BD was 8.14 (95% CI 71.74-90.54) days, whereas it increased to 98.15 (95% CI 85.67-110.63) days for people with an earlier diagnosis.

The survival time of patients with comorbid anxiety disorder diverged from those who did not have symptoms of it over time, with a cumulative risk of use discontinuation of 0.789 and 0.500 at 60 days, respectively. Relatively few patients continued to use the app after the very first month overall, but among those who were still using it, participants with anxiety disorder continued to show a survival disadvantage over those who did not experience it. The mean estimation of app use of the latter group was 95.07 (95% CI 85.43-104.71) days, whereas that of participants with an anxiety disorder was 78.19 (95% CI 66.56-89.83) days; therefore, having an anxiety disorder significantly influenced app use. Nevertheless, anxiety disorder was self-reported based on what users consider anxiety; hence, we tried to see if there was homogeneity between self-reports and treatment prescriptions at study initiation. Analysis showed that there is a statistically significant relationship between anxiolytics use and self-reported anxiety disorder (P<.001), and 59.6% (99/166) of the users who reported having an anxiety disorder did use them, whereas anxiolytics use decreased to 37.5% (84/224) among the participants who did not report an anxiety disorder. As antidepressant drugs can also be used to treat a number of other conditions, including anxiety disorders, we also analyzed association of self-reported anxiety disorder with anxiolytics, along with the most widely prescribed antidepressants for anxiety. It turned out that 74.7% (124/166) of the users who reported having an anxiety disorder did use these medications, despite the guideline recommendations to avoid antidepressants in BD.

Furthermore, we performed a Cox (proportional hazards) regression analysis to estimate the hazard of discontinuing app use for regular users, given their prognostic variables. The results of the Cox model analysis are presented in Table 5.

Table 5. Cox regression model analysis of user survival using the SIMPLe app 1.5[a].

Characteristics	Coefficient	Exp (coefficient; 95% CI)	*P* value	Concordance, mean (SE)
Age	–0.016	0.984 (0.971-0.998)	.02	0.589 (–0.019)
Country				N/A[b]
Spain	—[c]	—	—	
Chile	–0.143	0.867 (0.588-1.278)	.47	
Argentina	0.091	1.096 (0.748-1.605)	.63	
Mexico	–0.121	0.886 (0.495-1.585)	.68	
Colombia	0.318	1.375 (0.797-2.37)	.25	
Other	0.141	1.152 (0.779-1.703)	.47	
Anxiety disorder	Yes	0.233	1.262 (0.975-1.634)	.07
Antipsychotic	Yes	0.334	1.396 (1.058-1.843)	.02

[a]Schoenfeld residuals to check the proportional-hazards assumption: age (χ^2_1=0.006; P=.94), anxiety disorder (χ^2_1=2.7; P=.10), and antipsychotic (χ^2_1=0.1; P=.72).

[b]N/A: not applicable.

[c]Our Cox model analyzed the risk of discontinuing the app use that participants from different nationalities had in comparison with Spanish participants. This row was maintained in the table to make clear that Spain was not included in the category *Other*.

The variables age, anxiety disorder, antipsychotic, and country were explored (Table 5). In this model, age is suggested as a protective factor for app use discontinuation, that is, the older the individual, the lower the risk of discontinuing app use. The regression coefficient for age was –0.016 (P=.02), which would imply a better engagement for older individuals. We calculated the percentage change in hazard rate for years' increase in age using the formula $100 \times (e^{(-0.016 \times 10)} - 1) = -14.8$, which allows us to estimate that a user older by 10 years would have a 14.8% reduction in their hazard compared with a user younger by 10 years.

We did not observe statistically significant differences among countries. Participants from the countries analyzed did not have a significantly different risk of discontinuing app use compared with Spanish users.

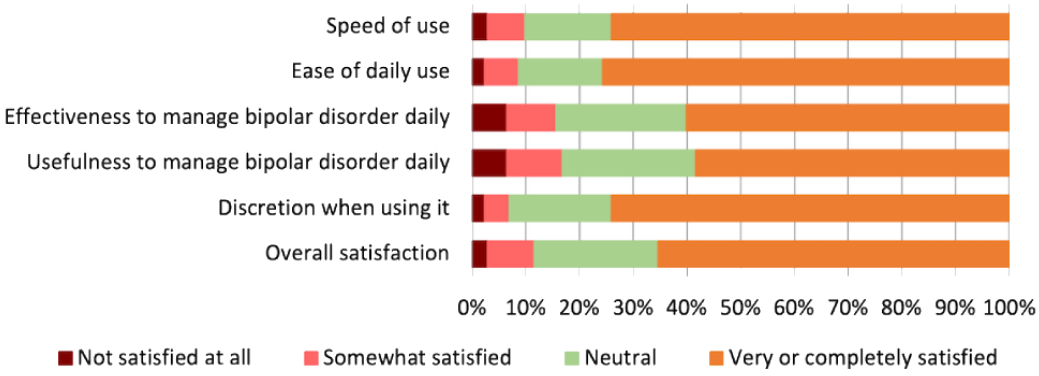
The regression coefficient for taking antipsychotics is statistically significant (coefficient=0.334, 95% CI 1.058-1.843; P=.02), which suggests that this variable is a risk factor and that users who take antipsychotics have a 33% hazard for discontinuing app use.

Usability, Satisfaction, and Perceived Usefulness

The analysis of user evaluation of the SIMPLe app contained in this section was exclusively performed with data of the 173 participants who used the app and completed the follow-up assessment too.

The mean raw SUS score was 77.05 (SD 17.21), which is above average at the 75th percentile. As shown in Figure 3, 74% (128/173) of the users were highly or completely satisfied with the speed and discretion of the app and 75.7% (131/173) of the users were satisfied with the ease of daily use. A high or complete level of overall satisfaction was rated by 65.3% (113/173) of the users.

Figure 3. Satisfaction with the SIMPLe app. The bars denote the percentage of satisfaction of users who responded to the follow-up questionnaire (n=173) after having experienced using the app.



With regard to usefulness, most users found somewhat useful or very useful the following features and functions: daily test (125/173, 72.3%), mood chart (127/173, 73.4%), personalized psychoeducational messages (117/173, 67.6%), weekly test (117/173, 67.6%), stressful events record (108/173, 62.4%), emergency alert notifications (103/173, 59.6%), the enabled option to share the mood chart (94/173, 54.3%), and the prodromes module (87/173, 50.3%).

Among the users who registered the reason for discontinuing using the app (91/173, 52.6%) by answering a multiple-choice question, 28.6% (26/91) found it very repetitive, 23.1% (21/91) had technical issues, 17.6% (16/91) did not find it useful, 16.5% (15/91) stated that it was an undesired daily reminder of their condition, 14.3% (13/91) stated that it affected smartphone performance, 13.2% (12/91) gave other reasons, 7.7% (7/91) were concerned about the stigma attached to having it installed on their smartphone, 4.4% (4/91) relapsed, and 3.3% (3/91) found it difficult to use.

Users suggested app improvements by responding to a multiple-choice question. The improvements more frequently suggested were enabling a personalized plan to follow when potential relapse symptoms appear (122/173, 70.5%), personalization of stressful events (103/173, 59.5%), and a wider variety of psychoeducational messages (99/173, 57.2%).

Discussion

After proving positive outcomes regarding satisfaction, usability, and helpfulness in previous research, the SIMPLe app was launched to the real world to make it available worldwide within the context of BD treatment.

Principal Findings

The outcomes of this real-world study represent the first attempt to evaluate, by means of survival analysis, use, retention patterns, and engagement of a large-scale wide-reaching app-based intervention providing psychoeducational content to patients with BD.

One of the most important advantages of the data collected through smartphone apps in clinical studies is the continuous and granular characteristics of the data registered at servers. Using detailed log and use data to examine predictive factors allows an understanding of the engagement and its underlying mechanisms. It will also aid optimization of smartphone-based interventions and improvement in the real-world uptake of self-management apps for BD, as well as in clinical benefits and associated outcomes [33].

Comparison With Prior Work

Survival analysis is not a new idea in statistics, and it is frequently used in several medical fields; in fact, it was considered the main outcome measure in the seminal works by Colom [34], who first showed the efficacy of group psychoeducation of patients in participants with BD. However, this method is a novel approach to quantify user retention and engagement of mobile apps that few studies have previously applied in nonpsychiatric populations [35,36]. Interestingly, Chien et al [37] used machine learning techniques to identify heterogeneity in patient engagement with internet-based cognitive behavioral therapy (CBT) for symptoms of depression and anxiety and found that patterns of patient behavior may reveal different modalities of engagement.

The aforementioned results are in line with one of the largest data sets on engagement in remote digital health compiled to date [38] (there is significant attrition in remote research). In addition, Pratap et al [38] observed indicators of retention in remote digital health studies through survival analysis, which revealed the factors associated with an increase in participant retention time, including older age (an increase of 4 days). In contrast, our study revealed that years since onset and years since diagnosis of BD (2 variables highly related to age) had a significant impact on app use; mobile app median use of participants with earlier disease onset and diagnosis increased by 32 days and 29 days compared with individuals with a more recent onset and more recent diagnosis, respectively. Similarly, when comparing potential predictors of traditional group psychoeducation in BD in a digital format, Reinares et al [39] identified that receiving an early diagnosis of BD may indicate a better response to face-to-face group psychoeducation. Other factors with a significant contribution to the risk of discontinuing app use (that are also highly associated with age) were housing and employment status. Retention duration of unemployed participants and those who shared a house was lower than that of users with other housing statuses. It is obvious that having a recent onset and diagnosis, sharing a house, and unemployment are more common in younger populations, suggesting a disadvantage regarding app use.

Our findings are consistent with previously collected preliminary use data on the SIMPLe app [22]. We previously observed that older age was a predictor significantly associated with higher odds of retention. In addition, the attrition rate of the program was still high, but this time our research focused on retention factors. Furthermore, overall satisfaction of the participants was quite positive because 65.3% (113/173) were highly or completely satisfied with SIMPLe in the context of low retention. The satisfaction, usability, and helpfulness outcomes are in line with our previous results deriving from an implementation feasibility study of SIMPLe. In terms of usefulness, the best-rated features and functions of the app were daily and weekly tests, mood charts, and personalized psychoeducational messages. The last-named is a unique and differentiating feature that other mobile-based platforms for BD do not offer [40]. On the basis of the information collected through screening tests on mood, energy, sleep time, medication adherence, and irritability, the SIMPLe app pushes daily pop-up notifications with a short psychoeducational message containing brief information (usually fewer than 50 words) on how to deal with specific situations to avoid relapses. Psychoeducational short messages carefully provided to cover detected specific user needs are the closest way to feature the *human touch*, and they fulfill quite convincingly our initial intention to partly replicate the successful Barcelona Bipolar Psychoeducation model [18] in an app so that it could de-escalate treatment costs and make combined therapy (psychoeducation and psychopharmacology) available to the greatest number of affected individuals.

In contrast, Faurholt-Jepsen et al [41] investigated the effect of smartphone-based patient-reported and objective monitoring, including a mood prediction system. The collected objective smartphone data included phone use, social activity, and mobility. Patients with BD were randomized to the Monsenso monitoring system (Monsenso A/S) or to standard treatment. Clinical feedback was established for patients in the intervention group in the case of signs of impairment (eg, lack of self-monitoring data). Overall, there was no effect of smartphone-based monitoring on symptoms compared with the control group. The intervention group adhered to the daily self-monitoring 72.6% (196/270) of the days over the 9-month study period.

Yet another project [42] evaluated 2 mobile phone–augmented interventions: an in-person session of CBT combined with automated thought-challenging and adaptive behavior delivered through mobile devices and an in-person session of psychoeducation with mobile interaction involving self-monitoring of symptoms. The retention rates were 77% for the self-monitoring group and 91% for the CBT condition at 6 months. Follow-up in both active conditions consisted of telephone calls by the study therapist to remind participants of assessment appointments and encourage adherence. However, these outcomes are difficult to extrapolate and compare because the cited Californian study included patients with schizophrenia as well as patients with BD. Previously, Depp et al [43] had explored a similar approach integrating a mobile device–delivered intervention linking patient-reported mood states with self-management strategies, preceded by face-to-face brief psychoeducation in BD. At 12 weeks, the retention reported was 93%; however, it was only operationalized based on the percentage of participants who returned the borrowed mobile devices of the study.

At odds with the aforementioned studies, the SIMPLe app was designed as an independent self-management method targeting relapse prevention. For research purposes, the study team helped (remotely) participants to install the app and log into the system, as well as provided a brief explanation. Users were provided with a telephone number to contact the research team for further assistance in case they experienced technical issues. The retention rate of the original SIMPLe study was 74% (36/51) after 3 months of app use [21].

The OpenSIMPLe study differs from the others in that it is the only modality that does not involve some contact with a person providing support or human interaction. It is reasonable to assume that the lower retention rates of the OpenSIMPLe study may have been influenced by the absence of human support in comparison with the other studies assessing smartphone-based platforms; the latter were more demanding in terms of time and staff resources. Besides the notifications systems recalling adherence in the aforementioned studies, the fact that participants established an alliance with clinical study staff and received previous face-to-face intervention or continuous telephone-delivered psychoeducation has influenced retention for certain [44].

It has been suggested that a positive relationship between app engagement and improvements in therapeutic outcomes in mental health and well-being may indicate the effectiveness of internet-based interventions [45]. Moreover, the type of engagement in terms of intervention use has demonstrated different relationships with outcomes [46]. However, there is a general tendency among mental health apps toward low retention and engagement rates [32,47]. At the same time, detailed engagement and retention rates are rarely reported in smartphone-based clinical trials but are of crucial importance to understand the underpinning of their effects. Those studies reporting and analyzing them frequently do so with heterogeneous and nonstandardized methodologies.

It is worth mentioning that we are comparing our retention, app use, and engagement data with other studies that used different parameters to measure these indicators for mental health apps and even used different criteria to assess them. Results from a systematic review conducted by Ng et al [32] indicated high variability across studies regarding the operationalization of the indicators of engagement, as well as highlighted the need to using objective criteria when assessing them. To date, the lack of clear consensus on the definition and standardization of these parameters represents a big obstacle to understanding feasibility and comparing results in the field of mobile-based apps and digital therapeutics. The real effectiveness of every intervention can be hardly estimated without quantifying the exact *dosage* of intervention that individuals receive, and engagement is not usually considered or it is evaluated erratically through effect size estimation in randomized controlled trials.

Limitations

Some limitations of this study should be noted, and caution should be exercised in generalizing results. First, our analyses relied on a rather heterogeneous sample, where participants differed in terms of sociodemographic, clinical, and psychological characteristics. This can be explained by the fact that we opened the platform to a real-world setting without considering inclusion and exclusion criteria that were too restrictive. However, the participants were a good representation of an unselected real-world population.

All measures were administered using exclusively self-reported web-based methods that did not allow us to get back in touch with participants who dropped out to collect feedback on the reasons for attrition. In addition, we did not have access to either medical records or passive data to validate the accuracy and reliability of the information provided, which may have influenced our sample and outcomes. As shown in Figure 1, there were only 2 participants excluded by the MDQ at screening. The MDQ is a popular, simple, and sensible screening instrument for the detection of BD. However, this tool is far from perfect [48]; it has low specificity, and it is likely that it did not discriminate among participants with a range of disorders such as borderline personality disorder. This is a common disadvantage in studies that screen participants through web-based methods exclusively.

A weakness of this study is that we limited use and retention analysis to the regular users of the SIMPLe 1.5 app and did not analyze other data from the occasional users, who did not use the app consistently. However, the aim of limiting these analyses to data provided by regular users was to avoid overestimation of use time and retention. In addition, sensitivity and homogeneity analysis confirmed that the data were coherent when we repeated the survival analysis with the whole sample of users; this showed that the selection of regular users in the study of app use prevented an overestimation of it, whereas the effect of the selection on survival probability is small.

Furthermore, only 44.4% (173/390) of the users completed the follow-up assessment, which implies some bias in the data collected regarding evaluation of the app because the variables measured at follow-up (including SUS, perceived usefulness, and satisfaction) were exclusively assessed by these users. An example that indicates this bias is that we found differences in the time range of app use between patients who used the SIMPLe app (n=390) and those who used it and completed the follow-up assessment (n=173); the former group used it a mean of 88.64 (SD 70.56) days, whereas this mean increased to 119.64 (SD 69.9) days in the latter group. However, the former group used the app slightly more every day on average (mean 0.92, SD 0.73) than the latter group (mean 0.83, SD 0.51). Therefore, these differences in terms of app use between the groups suggest that the users may have different profiles.

As this work was an exploratory study (ie, a flexible rather than structured approach to data collection was considered useful), there was no control group or alternative intervention for comparison of effects because the study was not designed to test the efficacy of the SIMPLe app. For the same reason and to avoid unlimited assessments that would probably result in the attrition rate soaring to unacceptable levels, we decided to keep control and covariate data to a minimum, which obviously represents at the same time something gained and something lost.

All the participants included came from Latin American or Spanish populations. The cultural characteristics of these origins may be difficult to generalize, but little is known about app adherence (or even drug or psychotherapy adherence) across cultures. This may become an exciting topic awaiting proper exploration.

Finally, it should be considered that the outcomes of this study deal with a high level of missing data derived from highly variable retention rates and lack of adherence after a few weeks of use among users of mental health apps, which is a common hindrance in internet-based research [12,49] that we tried to handle in an honest and rigorous manner.

Conclusions

The user retention rate of the app decreased at a rapid rate after each month until reaching only one-third of the users at 6 months. There exists a strong association between age and app engagement of individuals with BD. Other variables such as years lived with BD, diagnosis of an anxiety disorder, and taking antipsychotics seem to play a relevant role as well. We believe that an understanding of these associations will help clinicians in the definition of the most suitable user profiles for predicting trends of engagement, optimization of app prescription, and management.

Acknowledgments

The authors of this manuscript were indirectly supported by research grants from the Spanish Ministry of Economy and Competitiveness PI15/00588 and PI19/00009 (to FC) and FI20/00008 (to AGE); Instituto de Salud Carlos III, Subdirección General de Evaluación y Fomento de la Investigación; and Fondo Europeo de Desarrollo Regional, Unión Europea, *Una manera de hacer Europa*. Other sources of indirect support are a Juan Rodés (JR18/00021) contract from Instituto de Salud Carlos III (to DH) as well as Secretaria d'Universitats I Recerca del Departament d'Economia I Coneixement (2017 SGR 1365 to EV, August Pi i Sunyer Biomedical Research Institute, and 2017 SGR 134 to FC, Hospital del Mar Medical Research Institute). GA is supported by a Pons Bartran 2020 grant (PI046549). EV's research is supported by the Spanish Ministry of Science and Innovation integrated into the Plan Nacional de I+D+I and cofinanced by the Instituto de Salud Carlos III (Subdirección General de Evaluación) and the Fondo Europeo de Desarrollo Regional; the ISCIII; CIBER Mental Health; the Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement (2017 SGR 1365), the Centres de Recerca de Catalunya program, and the Departament de Salut de la Generalitat de Catalunya. The funding sources did not have any involvement in study design, collection, analysis,

interpretation of data, or the writing of this manuscript. The authors are very grateful to all the users of the SIMPLe app for their collaboration in this project. The technical development of the app software was commissioned to Sodep SA (Asunción, Paraguay).

Authors' Contributions

This study was conceived by FC, DH, and EV. DH and FC developed and maintained the project's website. FC, DH, and AGE were responsible for the methodology. JC assisted with data cleaning, statistical analysis, and interpretation of the results. Regarding the manuscript, the original draft was prepared by AGE, GA, and DH; review and editing were carried out by NAO, GA, EMM, FC, and VP; and supervision was by FC. EV and FC were responsible for funding acquisition. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

DH, EV, and FC designed the SIMPLe smartphone app mentioned in this study. The authors do not have any economic interests in the SIMPLe app, its use, or copyrights. EV has received grants and served as consultant, advisor or CME speaker for the following entities (unrelated to the present work): AB-Biotics, Abbott, Abbvie, Aimentia, Angelini, Biogen, Boehringer -Ingelheim, Casen-Recordati, Celon, Dainippon Sumitomo Pharma, Ferrer, Gedeon Richter, GH Research, Glaxo Smith-Kline, Janssen, Lundbeck, Organon, Otsuka, Sage, Sanofi-Aventis, Sunovion, Takeda, and Viatris. GA has received CME-related honoraria, or consulting fees from Janssen-Cilag, Lundbeck, and Angelini with no financial or other relationship relevant to the subject of this article.

References

1. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Dec 10;392(10159):1789-1858 [FREE Full text] [doi: 10.1016/S0140-6736(18)32279-7] [Medline: 30496104]
2. MacQueen G, Young LT, Joffe R. A review of psychosocial outcome in patients with bipolar disorder. Acta Psychiatr Scand 2001 Mar;103(3):163-170 [FREE Full text] [doi: 10.1034/j.1600-0447.2001.00059.x] [Medline: 11240572]
3. Michalak EE, Yatham LN, Lam RW. Quality of life in bipolar disorder: a review of the literature. Health Qual Life Outcomes 2005 Nov 15;3:72 [FREE Full text] [doi: 10.1186/1477-7525-3-72] [Medline: 16288650]
4. Dutta R, Boydell J, Kennedy N, Van Os J, Fearon P, Murray RM. Suicide and other causes of mortality in bipolar disorder: a longitudinal study. Psychol Med 2007 Mar 12;37(6):839-847 [FREE Full text] [doi: 10.1017/s0033291707000347] [Medline: 17349107]
5. Pompili M, Gonda X, Serafini G, Innamorati M, Sher L, Amore M, et al. Epidemiology of suicide in bipolar disorders: a systematic review of the literature. Bipolar Disord 2013 Aug 12;15(5):457-490 [FREE Full text] [doi: 10.1111/bdi.12087] [Medline: 23755739]
6. National Collaborating Centre for Mental Health (UK). Bipolar disorder: the NICE guideline on the assessment and management of bipolar disorder in adults, children and young people in primary and secondary care. In: National Institute for Health and Care Excellence: Clinical Guidelines. London, UK: The British Psychological Society and The Royal College of Psychiatrists; 2014.
7. Scott J, Colom F, Popova E, Benabarre A, Cruz N, Valenti M, et al. Long-term mental health resource utilization and cost of care following group psychoeducation or unstructured group support for bipolar disorders: a cost-benefit analysis. J Clin Psychiatry 2009 Mar 10;70(3):378-386 [FREE Full text] [doi: 10.4088/jcp.08m04333] [Medline: 19284929]
8. Saxena S, Thornicroft G, Knapp M, Whiteford H. Resources for mental health: scarcity, inequity, and inefficiency. Lancet 2007 Sep;370(9590):878-889 [FREE Full text] [doi: 10.1016/s0140-6736(07)61239-2] [Medline: 17804062]
9. Christensen H, Griffiths K, Evans K. E-Mental Health in Australia?: Implications of the Internet and Related Technologies for Policy. Canberra: Commonwealth Dept. of Health and Ageing; 2002.
10. World Health Organization. Global Diffusion of eHealth: Making Universal Health Coverage Achievable. Report of the Third Global Survey on eHealth. Geneva: World Health Organization; 2016.
11. Berry N, Lobban F, Emsley R, Bucci S. Acceptability of Interventions Delivered Online and Through Mobile Phones for People Who Experience Severe Mental Health Problems: A Systematic Review. J Med Internet Res 2016 May 31;18(5):e121 [FREE Full text] [doi: 10.2196/jmir.5250] [Medline: 27245693]
12. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018 Aug 05;21(3):116-119 [FREE Full text] [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
13. Lui JH, Marcus DK, Barry CT. Evidence-based apps? A review of mental health mobile applications in a psychotherapy context. Prof Psychol Res Pract 2017 Jun;48(3):199-210. [doi: 10.1037/pro0000122]
14. Larsen ME, Nicholas J, Christensen H. A systematic assessment of smartphone tools for suicide prevention. PLoS One 2016 Apr 13;11(4):e0152285 [FREE Full text] [doi: 10.1371/journal.pone.0152285] [Medline: 27073900]

15. Nicholas J, Larsen ME, Proudfoot J, Christensen H. Mobile apps for bipolar disorder: a systematic review of features and content quality. J Med Internet Res 2015 Aug 17;17(8):e198 [FREE Full text] [doi: 10.2196/jmir.4581] [Medline: 26283290]
16. Meurk C, Leung J, Hall W, Head BW, Whiteford H. Establishing and governing e-mental health care in australia: a systematic review of challenges and a call for policy-focussed research. J Med Internet Res 2016 Jan 13;18(1):e10 [FREE Full text] [doi: 10.2196/jmir.4827] [Medline: 26764181]
17. Wozney L, Newton AS, Gehring ND, Bennett K, Huguet A, Hartling L, et al. Implementation of eMental health care: viewpoints from key informants from organizations and agencies with eHealth mandates. BMC Med Inform Decis Mak 2017 Jun 02;17(1):78 [FREE Full text] [doi: 10.1186/s12911-017-0474-9] [Medline: 28577543]
18. Colom F, Vieta E. Psychoeducation Manual for Bipolar Disorder. Cambridge, United Kingdom: Cambridge University Press; 2006.
19. Colom F, Vieta E, Sánchez-Moreno J, Palomino-Otiniano R, Reinares M, Goikolea JM, et al. Group psychoeducation for stabilised bipolar disorders: 5-year outcome of a randomised clinical trial. Br J Psychiatry 2009 Mar 02;194(3):260-265 [FREE Full text] [doi: 10.1192/bjp.bp.107.040485] [Medline: 19252157]
20. Hidalgo-Mazzei D, Mateu A, Reinares M, Undurraga J, Bonnín CM, Sánchez-Moreno J, et al. Self-monitoring and psychoeducation in bipolar patients with a smart-phone application (SIMPLe) project: design, development and studies protocols. BMC Psychiatry 2015 Mar 20;15(1):52 [FREE Full text] [doi: 10.1186/s12888-015-0437-6] [Medline: 25884824]
21. Hidalgo-Mazzei D, Mateu A, Reinares M, Murru A, del Mar Bonnín C, Varo C, et al. Psychoeducation in bipolar disorder with a SIMPLe smartphone application: feasibility, acceptability and satisfaction. J Affect Disord 2016 Aug;200:58-66 [FREE Full text] [doi: 10.1016/j.jad.2016.04.042] [Medline: 27128358]
22. Hidalgo-Mazzei D, Reinares M, Mateu A, Nikolova VL, Bonnín CM, Samalin L, et al. OpenSIMPLe: A real-world implementation feasibility study of a smartphone-based psychoeducation programme for bipolar disorder. J Affect Disord 2018 Dec 01;241:436-445 [FREE Full text] [doi: 10.1016/j.jad.2018.08.048] [Medline: 30145515]
23. Hidalgo-Mazzei D, Reinares M, Mateu A, Juruena MF, Young AH, Pérez-Sola V, et al. Is a SIMPLe smartphone application capable of improving biological rhythms in bipolar disorder? J Affect Disord 2017 Dec 01;223:10-16 [FREE Full text] [doi: 10.1016/j.jad.2017.07.028] [Medline: 28711743]
24. Levin JB, Krivenko A, Howland M, Schlachet R, Sajatovic M. Medication adherence in patients with bipolar disorder: a comprehensive review. CNS Drugs 2016 Sep 19;30(9):819-835 [FREE Full text] [doi: 10.1007/s40263-016-0368-x] [Medline: 27435356]
25. Proyecto SIMPLe. Bipolar Disorders Program Barcelona. URL: https://simplebipolarproject.org/ [accessed 2021-06-22]
26. Hirschfeld RM, Holzer C, Calabrese JR, Weissman M, Reed M, Davies M, et al. Validity of the mood disorder questionnaire: a general population study. Am J Psychiatry 2003 Jan;160(1):178-180 [FREE Full text] [doi: 10.1176/appi.ajp.160.1.178] [Medline: 12505821]
27. Vieta E, Sánchez-Moreno J, Bulbena A, Chamorro L, Ramos J, Artal J, EDHIPO (Hypomania Detection Study) Group. Cross validation with the mood disorder questionnaire (MDQ) of an instrument for the detection of hypomania in Spanish: the 32 item hypomania symptom check list (HCL-32). J Affect Disord 2007 Aug;101(1-3):43-55 [FREE Full text] [doi: 10.1016/j.jad.2006.09.040] [Medline: 17189651]
28. Bonnín CM, Yatham L, Michalak E, Martínez-Arán A, Dhanoa T, Torres I, et al. Psychometric properties of the well-being index (WHO-5) Spanish version in a sample of euthymic patients with bipolar disorder. J Affect Disord 2018 Mar 01;228:153-159 [FREE Full text] [doi: 10.1016/j.jad.2017.12.006] [Medline: 29248821]
29. Brooke J. Smart phone applications for people with brain injury: the project SUS - a quick and dirty usability scale usability and context. In: Usability Evaluation In Industry. London: CRC Press; 1996.
30. 5 ways to interpret a SUS score. Measuring U. URL: https://measuringu.com/interpret-sus-score/ [accessed 2021-05-05]
31. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123 [FREE Full text] [doi: 10.5555/2835587.2835589]
32. Ng MM, Firth J, Minen M, Torous J. User engagement in mental health apps: a review of measurement, reporting, and validity. Pyschiatr Serv 2019 Jul 01;70(7):538-544 [FREE Full text] [doi: 10.1176/appi.ps.201800519] [Medline: 30914003]
33. Scherer EA, Ben-Zeev D, Li Z, Kane JM. Analyzing mHealth engagement: joint models for intensively collected user engagement data. JMIR Mhealth Uhealth 2017 Jan 12;5(1):e1 [FREE Full text] [doi: 10.2196/mhealth.6474] [Medline: 28082257]
34. Colom F, Vieta E, Reinares M, Martínez-Arán A, Torrent C, Goikolea JM, et al. Psychoeducation efficacy in bipolar disorders: beyond compliance enhancement. J Clin Psychiatry 2003 Sep 15;64(9):1101-1105 [FREE Full text] [doi: 10.4088/jcp.v64n0917] [Medline: 14628987]
35. Edney S, Ryan JC, Olds T, Monroe C, Fraysse F, Vandelanotte C, et al. User engagement and attrition in an app-based physical activity intervention: secondary analysis of a randomized controlled trial. J Med Internet Res 2019 Nov 27;21(11):e14645 [FREE Full text] [doi: 10.2196/14645] [Medline: 31774402]
36. Lin Y, Chen S, Lin P, Tai A, Pan Y, Hsieh C, et al. Assessing user retention of a mobile app: survival analysis. JMIR Mhealth Uhealth 2020 Nov 26;8(11):e16309 [FREE Full text] [doi: 10.2196/16309] [Medline: 33242023]

37. Chien I, Enrique A, Palacios J, Regan T, Keegan D, Carter D, et al. A machine learning approach to understanding patterns of engagement with internet-delivered mental health interventions. JAMA Netw Open 2020 Jul 01;3(7):e2010791 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.10791] [Medline: 32678450]
38. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. NPJ Digit Med 2020 Feb 17;3(1):21 [FREE Full text] [doi: 10.1038/s41746-020-0224-8] [Medline: 32128451]
39. Reinares M, Pacchiarotti I, Solé B, García-Estela A, Rosa AR, Bonnín CM, et al. A prospective longitudinal study searching for predictors of response to group psychoeducation in bipolar disorder. J Affect Disord 2020 Sep 01;274:1113-1121 [FREE Full text] [doi: 10.1016/j.jad.2020.02.047] [Medline: 32663939]
40. Lagan S, Ramakrishnan A, Lamont E, Ramakrishnan A, Frye M, Torous J. Digital health developments and drawbacks: a review and analysis of top-returned apps for bipolar disorder. Int J Bipolar Disord 2020 Dec 01;8(1):39 [FREE Full text] [doi: 10.1186/s40345-020-00202-4] [Medline: 33259047]
41. Faurholt-Jepsen M, Frost M, Christensen EM, Bardram JE, Vinberg M, Kessing LV. The effect of smartphone-based monitoring on illness activity in bipolar disorder: the MONARCA II randomized controlled single-blinded trial. Psychol Med 2019 Apr 04;50(5):838-848 [FREE Full text] [doi: 10.1017/S0033291719000710] [Medline: 30944054]
42. Depp C, Perivoliotis D, Holden J, Dorr J, Granholm E. Single-session mobile-augmented intervention in serious mental illness: a three-arm randomized controlled trial. Schizophr Bull 2019 Jun 18;45(4):752-762 [FREE Full text] [doi: 10.1093/schbul/sby135] [Medline: 30281086]
43. Depp CA, Ceglowski J, Wang VC, Yaghouti F, Mausbach BT, Thompson WK, et al. Augmenting psychoeducation with a mobile intervention for bipolar disorder: a randomized controlled trial. J Affect Disord 2015 Mar 15;174:23-30 [FREE Full text] [doi: 10.1016/j.jad.2014.10.053] [Medline: 25479050]
44. Baumel A, Kane JM. Examining predictors of real-world user engagement with self-guided ehealth interventions: analysis of mobile apps and websites using a novel dataset. J Med Internet Res 2018 Dec 14;20(12):e11491 [FREE Full text] [doi: 10.2196/11491] [Medline: 30552077]
45. Carpenter J, Crutchley P, Zilca RD, Schwartz HA, Smith LK, Cobb AM, et al. Seeing the "Big" picture: big data methods for exploring relationships between usage, language, and outcome in internet intervention data. J Med Internet Res 2016 Aug 31;18(8):e241 [FREE Full text] [doi: 10.2196/jmir.5725] [Medline: 27580524]
46. Short CE, DeSmet A, Woods C, Williams SL, Maher C, Middelweerd A, et al. Measuring engagement in ehealth and mhealth behavior change interventions: viewpoint of methodologies. J Med Internet Res 2018 Nov 16;20(11):e292 [FREE Full text] [doi: 10.2196/jmir.9397] [Medline: 30446482]
47. Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. J Med Internet Res 2019 Sep 25;21(9):e14567 [FREE Full text] [doi: 10.2196/14567] [Medline: 31573916]
48. Zimmerman M, Galione JN, Ruggero CJ, Chelminski I, Dalrymple K, Young D. Are screening scales for bipolar disorder good enough to be used in clinical practice? Compr Psychiatry 2011 Nov;52(6):600-606 [FREE Full text] [doi: 10.1016/j.comppsych.2011.01.004] [Medline: 21406301]
49. Arean PA, Hallgren KA, Jordan JT, Gazzaley A, Atkins DC, Heagerty PJ, et al. The use and effectiveness of mobile apps for depression: results from a fully remote clinical trial. J Med Internet Res 2016 Dec 20;18(12):e330 [FREE Full text] [doi: 10.2196/jmir.6482] [Medline: 27998876]

Abbreviations

BD: bipolar disorder
CBT: cognitive behavioral therapy
MDQ: Mood Disorder Questionnaire
SUS: System Usability Scale

Edited by R Kukafka; submitted 25.06.21; peer-reviewed by M Elhadi, E Chan, E Morton; comments to author 23.07.21; revised version received 16.09.21; accepted 29.10.21; published 02.02.22.

Please cite as:
García-Estela A, Cantillo J, Angarita-Osorio N, Mur-Milà E, Anmella G, Pérez V, Vieta E, Hidalgo-Mazzei D, Colom F
Real-world Implementation of a Smartphone-Based Psychoeducation Program for Bipolar Disorder: Observational Ecological Study
J Med Internet Res 2022;24(2):e31565
URL: https://www.jmir.org/2022/2/e31565
doi:10.2196/31565
PMID:35107440

Original Paper

Prevalence, Factors, and Association of Electronic Communication Use With Patient-Perceived Quality of Care From the 2019 Health Information National Trends Survey 5-Cycle 3: Exploratory Study

Rumei Yang[1*], PhD; Kai Zeng[2*], MSc; Yun Jiang[3*], PhD

[1]School of Nursing, Nanjing Medical University, Nanjing, China

[2]School of Nursing, Southern Medical University, Guangzhou, China

[3]School of Nursing, University of Michigan, Ann Arbor, MI, United States

[*]all authors contributed equally

Corresponding Author:
Rumei Yang, PhD
School of Nursing
Nanjing Medical University
818 Tianyuan E Rd
Nanjing, 211166
China
Phone: 86 02586869558
Email: rumeiyang@njmu.edu.cn

Abstract

Background: Electronic communication (e-communication), referring to communication through electronic platforms such as the web, patient portal, or mobile phone, has become increasingly important, as it extends traditional in-person communication with fewer limitations of timing and locations. However, little is known about the current status of patients' use of e-communication with clinicians and whether the use is related to the better patient-perceived quality of care at the population level.

Objective: The aim of this study was to explore the prevalence of and the factors associated with e-communication use and the association of e-communication use with patient-perceived quality of care by using the nationally representative sample of the 2019 Health Information National Trends Survey 5 (HINTS 5)-Cycle 3.

Methods: Data from 5438 adult responders (mean age 49.04 years, range 18-98 years) were included in this analysis. Multiple logistic and linear regressions were conducted to explore responders' personal characteristics related to their use of e-communication with clinicians in the past 12 months and how their use was related to perceived quality of care. Descriptive analyses for e-communication use according to age groups were also performed. All analyses considered the complex survey design using the jackknife replication method.

Results: The overall prevalence of e-communication use was 60.3%, with a significantly lower prevalence in older adults (16.6%) than that in <45-year-old adults (41%) and 45-65-year-old adults (42.4%). All percentages are weighted; therefore, absolute values are not shown. American adults who used e-communication were more likely to be high school graduates (odds ratio [OR] 1.95, 95% CI 1.14-3.34; *P*=.02), some college degree holders (OR 3.34, 95% CI 1.84-6.05; *P*<.001), and college graduates or more (OR 4.89, 95% CI 2.67-8.95; *P*<.001). Further, people who were females (OR 1.47, 95% CI 1.18-1.82; *P*=.001), with a household income ≥US $50,000 (OR 1.63, 95% CI 1.23-2.16; *P*=.001), with more comorbidities (OR 1.22, 95% CI 1.07-1.40; *P*=.004), or having a regular health care provider (OR 2.62, 95% CI 1.98-3.47; *P*<.001), were more likely to use e-communication. In contrast, those who resided in rural areas (OR 0.61, 95% CI 0.43-0.88; *P*=.009) were less likely to use e-communication. After controlling for the sociodemographics, the number of comorbidities, and relationship factors (ie, having a regular provider and trusting a doctor), e-communication use was found to be significantly associated with better perceived quality of care (β=.12, 95% CI 0.02-0.22; *P*=.02).

Conclusions: This study confirmed the positive association between e-communication use and patient-perceived quality of care and suggested that policy-level attention should be raised to engage the socially disadvantaged (ie, those with lower levels of education and income, without a regular health care provider, and living in rural areas) to maximize e-communication use and to support better patient-perceived quality of care among American adults.

(J Med Internet Res 2022;24(2):e27167) doi:10.2196/27167

KEYWORDS

electronic communication; quality of care; person-related characteristics; patient preference; HINTS

Introduction

Effective patient-clinician communication is a critical component of high-quality patient-centered care. With the rapid diffusion of advanced technology, the use of electronic services such as email, text messaging, and patient portals as a platform of communication (ie, electronic communication [e-communication]) between patients and clinicians has become increasingly popular [1]. Evidence shows that patients are enthusiastic about e-communication with clinicians regarding a wide variety of clinical contexts such as chronic condition self-management and follow-up examinations [2-4]. e-Communication has become a valuable supplement to traditional in-person communication through office visits [5,6]. It has fundamentally improved patients' interactions with the health care system and their engagement in shared decision-making with clinicians [7,8].

Despite the increasing popularity and potential impacts of e-communication on health care services, the actual use of e-communication among various patient populations still remains relatively low [9-12]. A review of patient portals for adults with diabetes found that 29%-46% of adults registered an account, but only 27%-76% of them actually accessed the portal [12]. A study of an encrypted message system in a pediatric clinic showed that only 4.3% of parents of chronically ill children made use of the system [11]. Similarly, a study of Health Information National Trends Survey 5 (HINTS 5)-Cycle 3 data in 2003-2005 indicated that only 10% of adult internet users communicated with the clinicians through web-based communication services [9]. Age can be a potential factor affecting the use of e-communication [13,14]. Clarke et al's [14] study showed that young adults preferred text messaging, middle-aged adults preferred phone calls, and older adults preferred paper-based and in-person interactions with clinicians. These findings imply that the prevalence of e-communication use might be lower among older adults as compared to that among young and middle-aged adults. Considering older adults' needs for technology-enabled health care support can help them become the major users of e-communication. In recent years, older adults' adoption of information and communication technology has been increasing, and they are likely to increasingly incorporate digital technology into their daily life [15]. Given the ever evolving technology and various populations' needs for support, it is important to understand how e-communication use varies across different age groups. Another commonly reported factor associated with patients' use of e-communication is patient-clinician relationships [16-20], for example, how much one trusts information from a doctor can influence the person's decision-making for using e-communication [19].

All these barriers can presumably affect both patients' use of e-communication [20] and their perceptions of quality of care [21]. However, there lacks empirical evidence to quantify the association between e-communication use and patient-perceived quality of care [22,23]. Patient-perceived quality of care refers to patients' perception of health care services received based on their experiences of what actually happened during the care process [24]. As one of the essential indicators of care quality, patients' perception of quality of care is an important driver of patient satisfaction, reflecting their desire for individualized high-quality care, which is also the main goal for those providing the care [25]. Factors that affect patient-perceived quality of care mainly include person-related conditions such as the patients' age, sex, education level, and self-reported health status, and external objective care conditions such as the organizational structure of care, competence of health care personnel, the size of the hospital, inpatient stay and occupancy, and comfortable environment [26]. Patient-clinician communication has been reported as one of the major factors driving patient perception of quality care in addition to hospital staff responsiveness, the care transition process, and hospital environment [27]. In the era of digital health, particularly with the increased popularity of e-communication between patients and clinicians and extended health care efficiency, the use of e-communication may increase the patient-perceived quality of care as opposed to no use of any e-communication [21]. However, the lack of empirical evidence to quantify the effectiveness of e-communication on patient-perceived quality of care may delay the promotion of e-communication adoption and the development of new models of patient-clinician interaction to satisfy patients' needs for high-quality health care services [21].

The purposes of this study were to examine the prevalence of patients' use of e-communication with clinicians and the potential factors (in particular, person-related factors such as age) associated with their use of e-communication and to explore the potential association between e-communication use and patient-perceived quality of care. Based on previous literature reports [21,24,25,27], we hypothesized that patients' use of e-communication was related to better patient-perceived quality of care.

Methods

Data Source

Data used in this study were from the HINTS 5-Cycle 3 [28]. HINTS is a nationally representative survey designed to understand American adults' knowledge of, attitudes toward, and use of cancer- and health-related information [29]. HINTS 5-Cycle 3 used a single-mode mail survey, with a 2-stage sample design, including a stratified sample of addresses and a selected adult within each sampled household [28]. The data were collected from 5438 respondents from January to May 2019 (English version only), with an overall 30.3% response rate [28]. Comprehensive reports on the sampling design for the HINTS survey have been published elsewhere [28-30]. The

survey data were deidentified and are publicly available; institutional review board approval was not applicable.

Variables

Perceived Quality of Care

The outcome variable patient-perceived quality of care was assessed via self-report on a single question asking "*overall, how would you rate the quality of health care you received in the past 12 months?*" with a 5-point Likert scale from 1=poor to 5=excellent, with a high score indicating better perceived quality of care (see Table 1).

Table 1. Variables and survey measurements.

Variable	Survey measurement
Patient-perceived quality of care	Overall, how would you rate the quality of health care you received in the past 12 months? (1=poor to 5=excellent)
Use of electronic communication	
1	In the past 12 months, have you used a computer, smartphone, or other electronic means to communicate with a doctor or a doctor's office? (1=yes, 0=no)
2	Have you sent a text message to or received a text message from a doctor or other health care professional within the last 12 months? (1=yes, 0=no)
3	In the past 12 months, have you used your online medical record to securely message health care provider and staff (eg, email)? (1=yes, 0=no)
4	In the past 12 months, have you used your online medical record to add health information to share with your health care provider, such as health concerns, symptoms, and side effects? (1=yes, 0=no)
5	Have you shared health information from either an electronic monitoring device or smartphone with a health professional within the last 12 months? (1=yes, 0=no)
6	Have you electronically sent your medical information to another health care clinician? (1=yes, 0=no).
Sociodemographics	
1	Age (young: ≥18 and <45 years, middle-aged: ≥45 and <65 years, and older adults ≥65 years)
2	Sex (0=male, 1=female)
3	Education level (0=less than high school, 1=high school graduate, 2=some college, 3=college graduate or more)
4	Marital status (0=not married, 1=married or partnered)
5	Race/ethnicity (0=White, 1=African American, 2=Hispanic, 3=other)
6	Household income (0=<US $50,000; 1=≥US $50,000)
7	Living status (0=living with others, 1=living alone)
8	Residency (0=nonrural, 1=rural)
Comorbidities	The number of comorbidities: Has a doctor or other health professional ever told you that you had any of the following medical conditions? Choices for this question included cancer, hypertension, diabetes, heart condition, chronic lung disease, and depression, and a sum score was used.
Patient-clinician relationship	
1	Having a regular health care provider: Not including psychiatrists and other mental health professionals, is there a particular doctor, nurse, or other health professional that you see most often? (0=no, 1=yes)
2	Trusting a doctor: In general, how much would you trust information about health or medical topics from a doctor? (1=not at all to 4=a lot)

Use of e-Communication

Patients' use of e-communication with clinicians in the past 12 months, such as using the computer, smartphone, text messaging, web-based messaging, web-based medical records, or any other electronic means to share medical information, were assessed through 6 survey questions (see Table 1). Survey responders who answered "yes" to either of the 6 questions were considered having e-communication with their clinicians, defined as users, while responders who answered "no" to all 6 questions were considered as nonusers.

Sociodemographics and Comorbidities

Age was measured as a continuous variable in the HINTS 5-Cycle 3 and was categorized into 3 groups: young adults (≥18 and <45 years of age, 38.4%), middle-aged adults (≥45 and <65 years of age, 39.7%), and older adults (≥65 years of age, 19.7%). All percentages are weighted; therefore, absolute values are not shown. Other sociodemographic covariates included sex, education level, marital status, race/ethnicity, household income, living status, and residency. The number of comorbidities was a sum score of 6 doctor-diagnosed chronic conditions, namely,

cancer, diabetes, hypertension, heart disease, lung disease, and depression (see Table 1).

Patient-Clinician Relationship

Patient-clinician relationship variables included (1) having a regular health care provider (yes/no) and (2) trusting a doctor (rating from 1=not at all to 4=a lot) (see Table 1).

Statistical Analysis

All analyses considered the complex survey design of the HINTS 5-Cycle 3 sample by using the HINTS-supplied final weights to estimate population estimates and 50 replicate weights to compute the standard errors with the jackknife replication approach [29]. Specifically, descriptive statistics were used to describe the prevalence and the characteristics of e-communication users and nonusers. Multiple logistic regression analyses were used to assess the association of sociodemographics and comorbidities (Model 1) and sociodemographics, comorbidities, plus patient-clinician relationship factors (Model 2) with e-communication use. Multiple linear regression analyses were used to examine the association between e-communication use and patient-perceived quality of care with the control of sociodemographics and comorbidities (Model 3) and the control of sociodemographics, comorbidities, plus patient-clinician relationship factors (Model 4). Missing data pattern analysis indicated that most variables had missing data <5% (see Table S1 in Multimedia Appendix 1). Multiple imputation was performed, and the pooled results of model 3 and model 4 were based on 50 imputed data sets using multiple imputation by chained equations. All analyses were conducted using Stata software (version 14; StataCorp). Results were reported as weighted point estimates and 95% CIs. The level of significance was .05.

Results

Prevalence and Characteristics of e-Communication Users

The overall prevalence of the use of e-communication was 60.3%. Most American adults who used e-communication with clinicians in the past 12 months were younger than 65 years, as older adults only accounted for 16.6% of e-communication users but 25.7% of nonusers (see Table 2). Table 2 also displays that most e-communication users were females (53.9%), had at least some college (41.7%), and 36.4% college graduates or more, were White people (65%), currently married (59.9%), with a household income ≥US $50,000 (63.9%), and did not live alone (85%) or in rural areas (89.5%). e-Communication users and nonusers were significantly different in all person-related characteristics. In addition, significantly more e-communication users had a regular health care provider than e-communication nonusers (72.9% vs 51.4%, respectively; P<.001) (see Table 2). All percentages are weighted; therefore, absolute values are not shown.

Table 2. Sociodemographic characteristics and comorbidities of electronic communication users versus nonusers.[a]

Characteristics	All users (N=5438)	Nonusers (n=2092)	Users (n=3337)	*P* value
Age (years), mean (SD)	49.58 (17.58)	50.52 (19.06)	48.06 (16.36)	*.005*
Comorbidities, mean (SD)	1.12 (1.15)	0.99 (1.14)	1.08 (1.13)	.14
Trusting a doctor, mean (SD)	3.67 (0.58)	3.56 (0.66)	3.66 (0.59)	*.001*
Patient-perceived quality of care, mean (SD)	3.96 (0.93)	3.84 (0.92)	4.01 (0.93)	*.002*
Age categories (% weighted)[b]				*<.001*
Young adults (<45 years)	38.4	36.5	41.0	
Middle-aged adults (45-64 years)	39.7	37.9	42.4	
Older adults (≥65 years)	19.7	25.7	16.6	
Gender (female) (% weighted)[c]	50.1	47.2	53.9	*.003*
Education level (% weighted)[d]				*<.001*
Less than high school	6.8	12.2	3.6	
High school graduate	22.8	31.3	18.3	
Some college	39.1	37.7	41.7	
College graduate or more	28.7	18.7	36.4	
Marital status (married or partnered, % weighted)[e]	54	49.1	59.9	*<.001*
Race/ethnicity (% weighted)[f]				*.002*
White	58	60.9	65	
African American	10.3	13.1	10.2	
Hispanic	15.4	20	14.9	
Other	7.7	6	9.9	
Household income (≥US $50,000) (% weighted)[g]	54.5	42.2	63.9	*<.001*
Living alone (% weighted)[h]	16.9	21.9	15	*<.001*
Residency (rural) (% weighted)[i]	13.3	17.4	10.5	*.001*
Having a regular health care provider (yes) (% weighted)[j]	63.3	51.4	72.9	*<.001*
Use of electronic communication (yes) (% weighted)[k]	60.3	—[l]	—	—

[a]Absolute values are not provided in this table because the percentages are weighted. The absolute values are summarized in the Multimedia Appendix 2. Significant *P* values are italicized.

[b]Age categories (0=young adults, 1=middle-aged adults, 2=older adults).

[c]Gender (0=male, 1=female).

[d]Education (0=less than high school, 1=high school graduate, 2=some college, 3=college graduate or more).

[e]Marital status (0=not married, 1=married or partnered).

[f]Race/ethnicity (0= White, 1=African American, 2=Hispanic, 3=other).

[g]Household income (0=less than US $50,000, 1=≥US $50,000).

[h]Living alone (0=living with others, 1=living alone).

[i]Residency (0=nonrural, 1=rural).

[j]Having a regular health care provider (0=no, 1=yes).

[k]Use of electronic communication with a clinician (0=no, 1=yes).

[l]Not available.

Factors Associated With e-Communication

Table 3 presents the results of multiple logistic regression analyses on the sociodemographics, comorbidities, and patient-clinician relationship factors for e-communication use. In model 1, where only sociodemographic factors and comorbidities were considered, age (odds ratio [OR] 0.87, 95% CI 0.66-1.14), female (OR 1.44, 95% CI 1.17-1.77), education level (eg, for college graduates or more, OR 4.78, 95% CI 2.63-8.68), household income (OR 1.77, 95% CI 1.34-2.34), rural residency (OR 0.62, 95% CI 0.44-0.87), and number of comorbidities (OR 1.33, 95% CI 1.16-1.52) were associated with e-communication use (see Table 3). In model 2, after adding the relationship factors to the model, people who were females (OR 1.47, 95% CI 1.18-1.82), high school graduates (OR 1.95, 95% CI 1.14-3.34), having some college (OR 3.34, 95% CI 1.84-6.05), and college graduates or more (OR 4.89, 95% CI 2.67-8.95), with a household income at or greater than US $50,000 (OR 1.63, 95% CI 1.23-2.16), with more comorbidities (OR 1.22, 95% CI 1.07-1.40), or having a regular health care provider (OR 2.62, 95% CI 1.98-3.47) were more likely to use e-communication, whereas those who were older adults (OR 0.42, 95% CI 0.31-0.57) or rural residents (OR 0.61, 95% CI 0.43-0.88) were less likely to use e-communication.

Table 3. Factors associated with electronic communication.

Variables	Model 1[a]		Model 2[b]	
	Odds ratio (95% CI)	*P* value[c]	Odds ratio (95% CI)	*P* value[c]
Age				
Young adults (<45 years)	Ref[d]	Ref	Ref	Ref
Middle-aged adults (45-64 years)	0.87 (0.66-1.14)	.31	0.86 (0.65-1.15)	.30
Older adults (≥65 years)	0.51 (0.39-0.68)	*<.001*	0.42 (0.31-0.57)	*<.001*
Female	1.44 (1.17-1.77)	*.001*	1.47 (1.18-1.82)	*.001*
Education level				
Less than high school	Ref	Ref	Ref	Ref
High school graduate	1.92 (1.09-3.39)	*.03*	1.95 (1.14-3.34)	*.02*
Some college	3.32 (1.82-6.07)	*<.001*	3.34 (1.84-6.05)	*<.001*
College graduate or more	4.78 (2.63-8.68)	*<.001*	4.89 (2.67-8.95)	*<.001*
Married or partnered	1.28 (0.97-1.69)	.08	1.26 (0.94-1.68)	.12
Race/ethnicity				
White	Ref	Ref	Ref	Ref
African American	0.93 (0.62-1.38)	.70	1.03 (0.68-1.57)	.89
Hispanic	0.86 (0.63-1.17)	.32	1.04 (0.76-1.41)	.81
Other	1.44 (0.93-2.21)	.10	1.55 (1.01-2.39)	.05
Household income (≥US $50,000)	1.77 (1.34-2.34)	*<.001*	1.63 (1.23-2.16)	*.001*
Living alone	0.94 (0.69-1.29)	.70	0.95 (0.68-1.31)	.74
Rural residency	0.62 (0.44-0.87)	*.008*	0.61 (0.43-0.88)	*.009*
Number of comorbidities	1.33 (1.16-1.52)	*<.001*	1.22 (1.07-1.40)	*.004*
Having a regular health care provider (yes)	—[e]	—	2.62 (1.98-3.47)	*<.001*
Trusting a doctor	—	—	1.14 (0.96-1.37)	.14

[a]Model 1 adjusted for sociodemographic factors (eg, age categories, gender, education, marital status, race/ethnicity) and comorbidities.

[b]Model 2 adjusted for sociodemographics, comorbidities, plus relationship factors (eg, having a regular health care provider, trust a doctor).

[c]Significant *P* values are italicized.

[d]Ref: reference value.

[e]Not available.

Associations Between e-Communication Use and Patient-Perceived Quality of Care

Table 4 displays the results of the association between e-communication use and patient-perceived quality of care among American adults. After controlling for sociodemographic factors (age, gender, education, income), comorbidities, and patient-clinician relationship factors (having a regular health care provider, trust a doctor), the use of e-communication was statistically associated with better quality of care (β=.12, 95% CI 0.02-0.22; see Model 4 in Table 4).

Table 4. Association between electronic communication and patient-perceived quality of care based on 50 imputed data sets using chained equations.

Variables	Model 3[a]			Model 4[b]		
	β	95% CI	*P* value[c]	β	95% CI	*P* value[c]
Use of electronic communication	.20	0.09 to 0.30	*<.001*	.12	0.02 to 0.22	*.02*
Age[d]						
Young adults	Ref[e]	Ref	Ref	Ref	Ref	Ref
Middle-aged adults	.06	–0.06 to 0.17	.36	.06	–0.05 to 0.17	.27
Older adults	.24	0.12 to 0.36	*<.001*	.17	0.05 to 0.28	*.005*
Female	.00	–0.10 to 0.09	.92	–.01	–0.10 to 0.08	.84
Education level						
Less than high school	Ref	Ref	Ref	Ref	Ref	Ref
High school graduate	.09	–0.30 to 0.12	.39	–.07	–0.26 to 0.12	.45
Some college	–.07	–0.27 to 0.14	.51	–.09	–0.28 to 0.10	.32
College graduate or more	.01	–0.20 to 0.22	.90	–.05	–0.24 to 0.14	.62
Married or partnered	.03	–0.11 to 0.17	.65	.04	–0.09 to 0.17	.53
Race/ethnicity						
White	Ref	Ref	Ref	Ref	Ref	Ref
African American	–.07	–0.24 to 0.09	.39	–.01	–0.17 to 0.14	.86
Hispanic	.00	–0.15 to 0.15	.96	.06	–0.08 to 0.20	.40
Other	–.25	–0.46 to –0.04	*.02*	–.23	–0.44 to –0.03	*.02*
Household income (≥US $50,000)	.11	–0.01 to 0.23	.08	.08	–0.04 to 0.19	.19
Living alone	.05	–0.11 to 0.21	.55	.07	–0.08 to 0.22	.38
Rural residency	–.02	–0.18 to 0.15	.82	–.03	–0.20 to 0.13	.70
Number of comorbidities	–.06	–0.11 to –0.00	*.03*	–.08	–0.12 to –0.03	*.001*
Having a regular health care provider	—[f]	—	—	.47	0.39 to 0.55	*<.001*
Trusting a doctor	—	—	—	.20	0.10 to 0.31	*<.001*

[a]Model 3 adjusted for sociodemographic factors (eg, age categories, gender, education, marital status, race/ethnicity) and comorbidities.

[b]Model 4 adjusted for sociodemographics, comorbidities, plus relationship factors (eg, having a regular health care provider, trusting a doctor).

[c]Significant *P* values are italicized.

[d]Age categories: young adults (<45 years), middle-aged adults (45-64 years), older adults (≥65 years).

[e]Ref: reference value.

[f]Not available.

Discussion

Principal Findings

This study examined the prevalence of and factors associated with e-communication use and the potential association between e-communication use and patient-perceived quality of care in a nationally representative sample of American adults. To the best of our knowledge, this study is the first to explore the association of e-communication use with patient-perceived quality of care at the population level. Several important findings emerged in this study.

First, the majority of American adults (60.3%) used some forms of e-communication with clinicians throughout 2019, which was significantly higher than the reported 7% in 2003, 10% in 2005 [9], and 31.5% in 2014 from the previous HINTS [31]. This finding indicates that e-communication use has become increasingly popular for adults to interact with their clinicians. The increased prevalence rate can be attributed to the increased availability and popularity of electronic health devices [32,33] and supportive policies (eg, promoting patient access to their electronic medical records) [34]. Although our data showed an overall growing trend in the use of e-communication, it is important to note that older adults' use of e-communication still remained relatively low, and this rate was not much improved from that in 2003 and 2005 [12]. Literature indicates that older adults usually prefer direct in-person interactions with their clinicians [12], while there are increasing reports about older adults' positive attitude toward e-communication and their preference for email and messaging communication with clinicians that is similar to that for younger adults [35,36]. Our

finding suggests that there is still a gap in the actual use of e-communication between older adults and young adults [31,37,38]. More studies are needed to explore the practical challenges that older adults may encounter in the use of e-communication. Older adults are potentially the major users of e-communication, considering their high level of health care needs. It is important to develop appropriate e-communication support for this population for their better health outcomes.

In addition to age, we also found that the use of e-communication varied by gender, education, income, and residency, indicating that individuals who are females, with higher education, higher income, and more comorbidities, or who reside in nonrural areas were more like to use e-communication with their clinicians. This finding is congruent with reports of the general adoption of eHealth in literature [39-42]. Consistent with our finding, the positive association between education and e-communication usage was reported in previous studies [31,39], which can be interpreted as individuals who have higher education might have more eHealth literacy skills and technological capabilities [43] to help them better use electronic forms of information [31,39]. However, Senft and Everson's recent study [44] reported that individuals who had lower levels of education and had negative care coordination experiences are more likely to use eHealth activities to communicate with clinicians [44], indicating that personal health care experiences can possibly interplay with education and thus influence the use of e-communication. However, it is unclear whether the limited use of e-communication among rural residents is related to lack of internet connectivity or awareness of e-communication services [45]. Additional studies can be conducted for further exploration.

Compared to those who did not use e-communication in the past year, in this study, e-communication users were more likely to have a regular health care provider and reported better trust in information from a doctor. However, trusting a doctor was not an independent predictor of e-communication use when having a regular health care provider was controlled for in the model. A previous qualitative study has indicated that a trusting relationship between patient-clinician is a significant contributor to better online patient-clinician interactions [20,46]. Even those who tend to frequently seek web-based health information are more willing to use the information provided by their trusted clinicians for their health decision-making [17,47]. Our findings suggested that patients with a regular health care provider had the greatest association with their use of e-communication. It is possible that patients who have a regular health care provider have already built a trusting relationship with their clinicians. Given the importance of trust in a provider in the patient-centered care process, future research directly examining possible confounding of this factor using longitudinal data is recommended.

Finally, it is not surprising that this study found that the use of e-communication was an independent predictor of patient-perceived quality of care. In 2001, the Institute of Medicine suggested that e-communication could improve the quality of care [48]. The previous literature review demonstrates that e-communication provides a convenient way of patient-clinician interaction, has a positive impact on patient satisfaction while saving time for patients and clinicians, and has the potential to extend health care efficiency [21,49]. The benefits and challenges of e-communication have been well addressed in the literature, while its benefits for the quality of care may not have been clearly quantified previously. The measures of quality of care can vary by the dimensions of care and care processes [50]. However, this study focused on the measure of the patient-perceived quality of care, which solely reflected patients' perceptions of health care services received based on their experiences of care [24]. It did not mean to measure any technical clinical quality, for example, cholesterol screening [51]. There is increasing interest in patient-reported measures, as experiences with care are more easily understood by patients. In addition, previous literature demonstrated that the measure of patient experiences of care was related to measures of the technical quality of care, which can serve as valid summary measures of hospital quality [52]. These study findings were based on the analysis of nationally representative survey data, which should be generalizable to all American adults. The positive association between the use of e-communication and perceived quality of care confirms that e-communication can serve as an important tool to improve patient satisfaction and their perceptions of quality of care. This finding is particularly significant and applicable in the current COVID-19 pandemic when traditional in-person communication is less feasible. It is expected that e-communication will continuously replace an adequate portion of traditional face-to-face encounters and has the potential to transform the health care system [21]. Future research can be conducted to explore the sustainable long-term effects of e-communication on patient-centered care outcomes.

Limitations

This study has a few limitations. First, data were mainly based on self-reports, which might have introduced recall bias. Second, the survey questions regarding the use of e-communication did not specify the frequency of use; therefore, they did not accurately reflect responders' experiences of using e-communication and might affect their perceptions of quality of care. Third, a binary measure of e-communication use (yes/no) was used, which might result in the loss of information or power. However, considering the conceptual overlaps across 6 questions about e-communication behaviors in the survey, a combined continuous assessment for the number of e-communication behaviors would be conceptually inaccurate. Fourth, the e-communication was between patients and clinicians. However, the survey only focused on the patient side and thus, it was not possible to know clinicians' perceptions of e-communication use. Finally, the results could be underestimated by potential reverse causality owing to the nature of the study design. The prevalence of e-communication use was higher in our study than that that reported in previous studies. The difference may also be due to varying measurement methods across studies. In our study, we used 6 questions to measure e-communication, which are more than that used in other studies. Different measures might affect comparisons of the prevalence of e-communication use across studies.

Conclusions

American adults' use of e-communication with clinicians has been significantly increased in the past decade, which may be due to increased patient needs and advanced support from technologies and policies. As a convenient way of patient-clinician interaction, the use of e-communication is significantly associated with patient-perceived quality of care. The findings of multiple factors associated with e-communication use and the positive association between e-communication use and patient-perceived quality of care suggest that policy-level attention is needed to engage the socially disadvantaged (ie, those with lower levels of education and income, without a regular health care provider, and living in rural areas) to maximize the use of e-communication and to support better patient-perceived quality of care among American adults.

Acknowledgments

This study was funded by the National Natural Science Foundation of China (grant 72004098) and Nanjing Medical University, Nanjing, China (JX10631803 and NMUR2020006).

Authors' Contributions

RY proposed the research questions, conducted data cleaning, data analysis, and manuscript writing. YJ proposed the research questions, was involved in the data analysis, data interpretation, and manuscript writing. KZ was involved in data interpretation, manuscript writing, and proofreading.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Missing data information.
[DOCX File , 18 KB - jmir_v24i2e27167_app1.docx]

Multimedia Appendix 2
Absolute value data for weighted percentages.
[DOCX File , 13 KB - jmir_v24i2e27167_app2.docx]

References

1. Hemmat M, Ayatollahi H, Maleki MR, Saghafi F. Future Research in Health Information Technology: A Review. Perspect Health Inf Manag 2017;14(Winter):1b [FREE Full text] [Medline: 28566991]
2. Kim S, Fadem S. Communication matters: Exploring older adults' current use of patient portals. Int J Med Inform 2018 Dec;120:126-136. [doi: 10.1016/j.ijmedinf.2018.10.004] [Medline: 30409337]
3. Voruganti T, Grunfeld E, Makuwaza T, Bender JL. Web-Based Tools for Text-Based Patient-Provider Communication in Chronic Conditions: Scoping Review. J Med Internet Res 2017 Oct 27;19(10):e366 [FREE Full text] [doi: 10.2196/jmir.7987] [Medline: 29079552]
4. Menachemi N, Prickett CT, Brooks RG. The use of physician-patient email: a follow-up examination of adoption and best-practice adherence 2005-2008. J Med Internet Res 2011 Feb 25;13(1):e23 [FREE Full text] [doi: 10.2196/jmir.1578] [Medline: 21447468]
5. Zhong X, Park J, Liang M, Shi F, Budd PR, Sprague JL, et al. Characteristics of Patients Using Different Patient Portal Functions and the Impact on Primary Care Service Utilization and Appointment Adherence: Retrospective Observational Study. J Med Internet Res 2020 Feb 25;22(2):e14410 [FREE Full text] [doi: 10.2196/14410] [Medline: 32130124]
6. Haun JN, Lind JD, Shimada SL, Martin TL, Gosline RM, Antinori N, et al. Evaluating user experiences of the secure messaging tool on the Veterans Affairs' patient portal system. J Med Internet Res 2014 Mar 06;16(3):e75 [FREE Full text] [doi: 10.2196/jmir.2976] [Medline: 24610454]
7. Baldwin FD. The doctor is in. With physician-patient messaging via Web portals, everyone is better served. Healthc Inform 2002 Oct;19(10):44-48. [Medline: 12440305]
8. Zhong X, Liang M, Sanchez R, Yu M, Budd PR, Sprague JL, et al. On the effect of electronic patient portal on primary care utilization and appointment adherence. BMC Med Inform Decis Mak 2018 Oct 16;18(1):84 [FREE Full text] [doi: 10.1186/s12911-018-0669-8] [Medline: 30326876]
9. Beckjord EB, Finney Rutten LJ, Squiers L, Arora NK, Volckmann L, Moser RP, et al. Use of the internet to communicate with health care providers in the United States: estimates from the 2003 and 2005 Health Information National Trends Surveys (HINTS). J Med Internet Res 2007 Jul 12;9(3):e20 [FREE Full text] [doi: 10.2196/jmir.9.3.e20] [Medline: 17627929]

10. Stamenova V, Agarwal P, Kelley L, Fujioka J, Nguyen M, Phung M, et al. Uptake and patient and provider communication modality preferences of virtual visits in primary care: a retrospective cohort study in Canada. BMJ Open 2020 Jul 06;10(7):e037064 [FREE Full text] [doi: 10.1136/bmjopen-2020-037064] [Medline: 32636284]
11. Hsiao AL, Bazzy-Asaad A, Tolomeo C, Edmonds D, Belton B, Benin AL. Secure web messaging in a pediatric chronic care clinic: a slow takeoff of "kids' airmail". Pediatrics 2011 Feb;127(2):e406-e413. [doi: 10.1542/peds.2010-1086] [Medline: 21220392]
12. Sun R, Korytkowski MT, Sereika SM, Saul MI, Li D, Burke LE. Patient Portal Use in Diabetes Management: Literature Review. JMIR Diabetes 2018 Nov 06;3(4):e11199 [FREE Full text] [doi: 10.2196/11199] [Medline: 30401665]
13. Abolhassani N, Santos-Eggimann B, Chiolero A, Santschi V, Henchoz Y. Readiness to accept health information and communication technologies: A population-based survey of community-dwelling older adults. Int J Med Inform 2019 Oct;130:103950. [doi: 10.1016/j.ijmedinf.2019.08.010] [Medline: 31446357]
14. Clarke MA, Fruhling AL, Sitorius M, Windle TA, Bernard TL, Windle JR. Impact of Age on Patients' Communication and Technology Preferences in the Era of Meaningful Use: Mixed Methods Study. J Med Internet Res 2020 Jun 01;22(6):e13470 [FREE Full text] [doi: 10.2196/13470] [Medline: 32478658]
15. Anderson M, Perrin A. Tech adoption climbs among older adults. Pew Research Center. 2017. URL: https://www.pewresearch.org/internet/2017/05/17/tech-adoption-climbs-among-older-adults/ [accessed 2021-06-20]
16. Ochoa A, Kitayama K, Uijtdehaage S, Vermillion M, Eaton M, Carpio F, et al. Patient and provider perspectives on the potential value and use of a bilingual online patient portal in a Spanish-speaking safety-net population. J Am Med Inform Assoc 2017 Nov 01;24(6):1160-1164 [FREE Full text] [doi: 10.1093/jamia/ocx040] [Medline: 28460130]
17. Zhao J, Song B, Anand E, Schwartz D, Panesar M, Jackson G, et al. Barriers, Facilitators, and Solutions to Optimal Patient Portal and Personal Health Record Use: A Systematic Review of the Literature. AMIA Annu Symp Proc 2017;2017:1913-1922 [FREE Full text] [Medline: 29854263]
18. Liederman EM, Lee JC, Baquero VH, Seites PG. Patient-physician web messaging. The impact on message volume and satisfaction. J Gen Intern Med 2005 Jan;20(1):52-57. [doi: 10.1111/j.1525-1497.2005.40009.x] [Medline: 15693928]
19. Lyles CR, Sarkar U, Ralston JD, Adler N, Schillinger D, Moffet HH, et al. Patient-provider communication and trust in relation to use of an online patient portal among diabetes patients: The Diabetes and Aging Study. J Am Med Inform Assoc 2013;20(6):1128-1131 [FREE Full text] [doi: 10.1136/amiajnl-2012-001567] [Medline: 23676243]
20. Jiang S. The Relationship between Face-to-Face and Online Patient-Provider Communication: Examining the Moderating Roles of Patient Trust and Patient Satisfaction. Health Commun 2020 Mar;35(3):341-349. [doi: 10.1080/10410236.2018.1563030] [Medline: 30821506]
21. Tabatabai S. Effects of physician-patient electronic communications on the quality of care. International Journal of Reliable and Quality E-Healthcare 2013 Apr 1 2013;2(2):54-62. [doi: 10.4018/ijrqeh.2013040105]
22. Walsh C, Siegler EL, Cheston E, O'Donnell H, Collins S, Stein D, Informatics Intervention Research Collaboration (I2RC). Provider-to-provider electronic communication in the era of meaningful use: a review of the evidence. J Hosp Med 2013 Oct;8(10):589-597 [FREE Full text] [doi: 10.1002/jhm.2082] [Medline: 24101544]
23. Alkureishi MA, Lee WW, Lyons M, Press VG, Imam S, Nkansah-Amankra A, et al. Impact of Electronic Medical Record Use on the Patient-Doctor Relationship and Communication: A Systematic Review. J Gen Intern Med 2016 May;31(5):548-560 [FREE Full text] [doi: 10.1007/s11606-015-3582-1] [Medline: 26786877]
24. Gishu T, Weldetsadik AY, Tekleab AM. Patients' perception of quality of nursing care; a tertiary center experience from Ethiopia. BMC Nurs 2019;18:37 [FREE Full text] [doi: 10.1186/s12912-019-0361-z] [Medline: 31427889]
25. Dansereau E, Masiye F, Gakidou E, Masters SH, Burstein R, Kumar S. Patient satisfaction and perceived quality of care: evidence from a cross-sectional national exit survey of HIV and non-HIV service users in Zambia. BMJ Open 2015 Dec 30;5(12):e009700 [FREE Full text] [doi: 10.1136/bmjopen-2015-009700] [Medline: 26719321]
26. Grøndahl VA, Kirchhoff JW, Lauvli Andersen K, Sørby LA, Andreassen HM, Skaug E, et al. Health care quality from the patients' perspective: A comparative study between an old and a new, high-tech hospital. JMDH 2018 Oct;Volume 11:591-600. [doi: 10.2147/jmdh.s176630]
27. Freudenberger DC, Baker EA, Siljander MP, Rohde RS. Factors Driving Patient Perception of Quality Care After Primary Total Hip and Total Knee Arthroplasty. JAAOS Glob Res Rev 2018;2(11):e061. [doi: 10.5435/jaaosglobal-d-18-00061]
28. Westat. Cycle 3 methodology report. Health Information National Trends Survey 5 (HINTS 5). 2019. URL: https://hints.cancer.gov/docs/Instruments/HINTS5_Cycle3_MethodologyReport.pdf [accessed 2021-06-20]
29. Nelson DE, Kreps GL, Hesse BW, Croyle RT, Willis G, Arora NK, et al. The Health Information National Trends Survey (HINTS): development, design, and dissemination. J Health Commun 2004;9(5):443-60; discussion 81. [doi: 10.1080/10810730490504233] [Medline: 15513791]
30. Turner K, Clary A, Hong Y, Alishahi Tabriz A, Shea CM. Patient Portal Barriers and Group Differences: Cross-Sectional National Survey Study. J Med Internet Res 2020 Sep 17;22(9):e18870 [FREE Full text] [doi: 10.2196/18870] [Medline: 32940620]
31. Spooner KK, Salemi JL, Salihu HM, Zoorob RJ. eHealth patient-provider communication in the United States: interest, inequalities, and predictors. J Am Med Inform Assoc 2017 Apr 01;24(e1):e18-e27 [FREE Full text] [doi: 10.1093/jamia/ocw087] [Medline: 27497797]

32. Tidwell LC, Walther JB. Computer-Mediated Communication Effects on Disclosure, Impressions, and Interpersonal Evaluations: Getting to Know One Another a Bit at a Time. Human Comm Res 2002 Jul;28(3):317-348. [doi: 10.1111/j.1468-2958.2002.tb00811.x]
33. Eveleigh RM, Muskens E, van Ravesteijn H, van Dijk I, van Rijswijk E, Lucassen P. An overview of 19 instruments assessing the doctor-patient relationship: different models or concepts are used. J Clin Epidemiol 2012 Jan;65(1):10-15. [doi: 10.1016/j.jclinepi.2011.05.011] [Medline: 22118265]
34. Cho AH, Arar NH, Edelman DE, Hartwell PH, Oddone EZ, Yancy WS. Do diabetic veterans use the Internet? Self-reported usage, skills, and interest in using My HealtheVet Web portal. Telemed J E Health 2010 Jun;16(5):595-602. [doi: 10.1089/tmj.2009.0164] [Medline: 20575727]
35. Lam R, Lin VS, Senelick WS, Tran H, Moore AA, Koretz B. Older adult consumers' attitudes and preferences on electronic patient-physician messaging. Am J Manag Care 2013 Nov;19(10 Spec No):eSP7-eS11 [FREE Full text] [Medline: 24511886]
36. Joe J, Demiris G. Older adults and mobile phones for health: a review. J Biomed Inform 2013 Oct;46(5):947-954 [FREE Full text] [doi: 10.1016/j.jbi.2013.06.008] [Medline: 23810858]
37. Saied A, Sherry SJ, Castricone DJ, Perry KM, Katz SC, Somasundar P. Age-related trends in utilization of the internet and electronic communication devices for coordination of cancer care in elderly patients. J Geriatr Oncol 2014 Apr;5(2):185-189. [doi: 10.1016/j.jgo.2013.11.001] [Medline: 24495698]
38. Kampmeijer R, Pavlova M, Tambor M, Golinowska S, Groot W. The use of e-health and m-health tools in health promotion and primary prevention among older adults: a systematic literature review. BMC Health Serv Res 2016 Sep 05;16 Suppl 5:290 [FREE Full text] [doi: 10.1186/s12913-016-1522-3] [Medline: 27608677]
39. Kontos E, Blake KD, Chou WS, Prestin A. Predictors of eHealth usage: insights on the digital divide from the Health Information National Trends Survey 2012. J Med Internet Res 2014 Jul 16;16(7):e172 [FREE Full text] [doi: 10.2196/jmir.3117] [Medline: 25048379]
40. Amo L. Education-Based Gaps in eHealth: A Weighted Logistic Regression Approach. J Med Internet Res 2016 Oct 12;18(10):e267 [FREE Full text] [doi: 10.2196/jmir.5188] [Medline: 27733329]
41. Nguyen A, Mosadeghi S, Almario CV. Persistent digital divide in access to and use of the Internet as a resource for health information: Results from a California population-based study. Int J Med Inform 2017 Jul;103:49-54. [doi: 10.1016/j.ijmedinf.2017.04.008] [Medline: 28551001]
42. Latulippe K, Hamel C, Giroux D. Social Health Inequalities and eHealth: A Literature Review With Qualitative Synthesis of Theoretical and Empirical Studies. J Med Internet Res 2017 Apr 27;19(4):e136 [FREE Full text] [doi: 10.2196/jmir.6731] [Medline: 28450271]
43. Vollbrecht H, Arora V, Otero S, Carey K, Meltzer D, Press VG. Evaluating the Need to Address Digital Literacy Among Hospitalized Patients: Cross-Sectional Observational Study. J Med Internet Res 2020 Jun 04;22(6):e17519 [FREE Full text] [doi: 10.2196/17519] [Medline: 32496196]
44. Senft N, Everson J. eHealth Engagement as a Response to Negative Healthcare Experiences: Cross-Sectional Survey Analysis. J Med Internet Res 2018 Dec 05;20(12):e11034 [FREE Full text] [doi: 10.2196/11034] [Medline: 30518513]
45. Greenberg AJ, Haney D, Blake KD, Moser RP, Hesse BW. Differences in Access to and Use of Electronic Personal Health Information Between Rural and Urban Residents in the United States. J Rural Health 2018 Feb;34 Suppl 1:s30-s38 [FREE Full text] [doi: 10.1111/jrh.12228] [Medline: 28075508]
46. Czaja R, Manfredi C, Price J. The determinants and consequences of information seeking among cancer patients. J Health Commun 2003;8(6):529-562. [doi: 10.1080/716100418] [Medline: 14690888]
47. Hu X, Bell RA, Kravitz RL, Orrange S. The prepared patient: information seeking of online support group members before their medical appointments. J Health Commun 2012;17(8):960-978. [doi: 10.1080/10810730.2011.650828] [Medline: 22574697]
48. Relman AS. Crossing the Quality Chasm: A New Health System for the 21st Century By the Committee on Quality of Health Care in America of the Institute of Medicine. N Engl J Med 2001 Aug 30;345(9):702-703. [doi: 10.1056/nejm200108303450917]
49. Bishop TF, Press MJ, Mendelsohn JL, Casalino LP. Electronic communication improves access, but barriers to its widespread adoption remain. Health Aff (Millwood) 2013 Aug;32(8):1361-1367 [FREE Full text] [doi: 10.1377/hlthaff.2012.1151] [Medline: 23918479]
50. Institute OM. Measuring the Quality of Health Care. Washington, DC: The National Academies Press; 1999.
51. Wilson IB, Landon BE, Marsden PV, Hirschhorn LR, McInnes K, Ding L, et al. Correlations among measures of quality in HIV care in the United States: cross sectional study. BMJ 2007 Nov 24;335(7629):1085. [doi: 10.1136/bmj.39364.520278.55]
52. Isaac T, Zaslavsky A, Cleary P, Landon B. The relationship between patients' perception of care and measures of hospital quality and safety. Health Serv Res 2010 Aug;45(4):1024-1040 [FREE Full text] [doi: 10.1111/j.1475-6773.2010.01122.x] [Medline: 20528990]

Abbreviations

e-Communication: electronic communication
HINTS: Health Information National Trends Survey
OR: odds ratio

Edited by R Kukafka, G Eysenbach; submitted 13.01.21; peer-reviewed by R Sun, N Senft, V Press, R Moser, A Hinson-Enslin, H Ayatollahi; comments to author 17.03.21; revised version received 25.06.21; accepted 10.11.21; published 04.02.22.

Please cite as:
Yang R, Zeng K, Jiang Y
Prevalence, Factors, and Association of Electronic Communication Use With Patient-Perceived Quality of Care From the 2019 Health Information National Trends Survey 5-Cycle 3: Exploratory Study
J Med Internet Res 2022;24(2):e27167
URL: https://www.jmir.org/2022/2/e27167
doi:10.2196/27167
PMID:35119369

Original Paper

Examining Diurnal Differences in Multidisciplinary Care Teams at a Pediatric Trauma Center Using Electronic Health Record Data: Social Network Analysis

Ashimiyu Durojaiye[1], MD, PhD; James Fackler[2], MD; Nicolette McGeorge[1], PhD; Kristen Webster[1], PhD; Hadi Kharrazi[3], MD, PhD; Ayse Gurses[1], MSc, MPH, PhD

[1]Armstrong Institute Center for Health Care Human Factors, Johns Hopkins University, Baltimore, MD, United States

[2]Division of Pediatric Anesthesiology and Critical Care Medicine, Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, MD, United States

[3]Division of Health Sciences Informatics, Johns Hopkins University School of Medicine, Baltimore, MD, United States

Corresponding Author:
Ayse Gurses, MSc, MPH, PhD
Armstrong Institute Center for Health Care Human Factors
Johns Hopkins University
750 E. Pratt St. 15th Floor
Baltimore, MD, 21202
United States
Phone: 1 410 637 4387
Email: agurses1@jhmi.edu

Abstract

Background: The care of pediatric trauma patients is delivered by multidisciplinary care teams with high fluidity that may vary in composition and organization depending on the time of day.

Objective: This study aims to identify and describe diurnal variations in multidisciplinary care teams taking care of pediatric trauma patients using social network analysis on electronic health record (EHR) data.

Methods: Metadata of clinical activities were extracted from the EHR and processed into an event log, which was divided into 6 different event logs based on shift (day or night) and location (emergency department, pediatric intensive care unit, and floor). Social networks were constructed from each event log by creating an edge among the functional roles captured within a similar time interval during a shift. Overlapping communities were identified from the social networks. Day and night network structures for each care location were compared and validated via comparison with secondary analysis of qualitatively derived care team data, obtained through semistructured interviews; and member-checking interviews with clinicians.

Results: There were 413 encounters in the 1-year study period, with 65.9% (272/413) and 34.1% (141/413) beginning during day and night shifts, respectively. A single community was identified at all locations during the day and in the pediatric intensive care unit at night, whereas multiple communities corresponding to individual specialty services were identified in the emergency department and on the floor at night. Members of the trauma service belonged to all communities, suggesting that they were responsible for care coordination. Health care professionals found the networks to be largely accurate representations of the composition of the care teams and the interactions among them.

Conclusions: Social network analysis was successfully used on EHR data to identify and describe diurnal differences in the composition and organization of multidisciplinary care teams at a pediatric trauma center.

(J Med Internet Res 2022;24(2):e30351) doi:10.2196/30351

KEYWORDS

pediatric trauma; multidisciplinary health team; multi-team systems; social network analysis; electronic health record; process mining; fluid teams

Introduction

Background

Multidisciplinary care teams in health care are increasingly being seen as a multi-team system (MTS) [1,2], where 2 or more teams communicate and coordinate to achieve overarching goals [3], such as providing optimal care. MTSs are different from traditional teams in that MTS constituent teams are interdependent, work across boundaries, share accountability, and function through a hierarchy of goals that determine how lower goals are accomplished to realize higher goals [3]. MTSs have three attributes as follows: (1) compositional attributes (eg, number of teams, size of teams, and changes in team composition), (2) linkage attributes (eg, interdependence, hierarchical structure, and communication structure), and (3) developmental attributes (eg, changes in team membership over time) [4], which support the specialization and flexibility that allow constituent teams to pursue lower goals while trying to achieve higher goals [5].

MTS are often seen in environments where tasks are ambiguous, multifaceted, dynamic, and urgent [5]. In health care, trauma teams that take care of patients with trauma are examples of MTS. The care of patients with trauma is complex, multidimensional, and time sensitive, requiring multidisciplinary collaboration among a variety of health care professionals (HCPs) with complementary expertise, [6] with high fluidity of team membership (ie, members join and others leave based on the needs of patients) [7]. In addition, staffing levels at trauma centers vary with the time of day and the day of week, such that services of HCPs deemed nonessential may not be available during *off hours* (nights and weekends) [8-11], necessitating changes and adaptation in MTS structures.

Assessment of MTS, as they perform their work in actual settings, is important to gain a better understanding of work as done (as opposed to work as imagined) [12] and to identify how to improve their performance given the *realities and variations of work* [1]. Social network analysis can enable the understanding and assessment of MTS at the compositional (ie, membership) and organizational (eg, subteam) levels [1,13]. Typically, such assessment is done through observation, which can be highly resource intensive and may not be practical to capture all the cognitive work of team members involved in the trauma. Moreover, a self-reported surveys [5], which relies exclusively on perceptions of care professionals may also be limited in its ability to provide rich details [3]. The ability to exploit *digital traces* [14], which may provide opportunities over survey data [15,16] or observational data, is desirable. Electronic health record (EHR) systems offer the opportunity to study the composition and organization of care teams working as part of an MTS [17]. EHRs capture many clinical activities that are performed by HCPs in the process of care delivery [17,18], and previous studies have shown the feasibility of obtaining plausible information about care teams from EHR data [17].

Objective

This study aims to identify MTS and demonstrate the dynamic nature of the compositional and organizational structures of the MTS by describing diurnal differences at various locations in a pediatric trauma center using EHR data.

Methods

Research Setting

This study was conducted as part of a larger research project (AHRQ R01HS023837) [19-21] aimed at redesigning pediatric trauma work systems based on health information to improve care transitions and patient safety. This study builds on a core methodology that has been previously described and validated [22,23]. The core methodology is reproduced here from data set subsection to “generation of master event log” subsection with necessary modifications for this paper.

Study Setting

This study was conducted at a large academic children’s medical hospital with a level I pediatric trauma center in the Eastern United States, which receives approximately 1000 pediatric patients with trauma per year. The participating hospital triages incoming patients into one of four trauma activation levels as follows: alpha (level I or highest severity), bravo (level II), critical trauma transfers (includes interfacility, but patients who are stable but critically injured, and is also known as a consult) and emergency department (ED) response, that are ordered by decreasing acuity and need for multidisciplinary care with ED response activations exclusively handled by the ED staff.

Trauma activation levels determine the composition of the trauma team, as specified by state [24] and institutional policy. The trauma team is derived from the ED staff, the general pediatric surgery service, pediatric intensive care unit (PICU), and the ancillary support staff (eg, child life specialists, chaplains, and social workers). Following resuscitation, if inpatient admission is required, patients with single-system injuries are admitted under the appropriate specialty service, whereas patients with multisystem injuries are admitted under the general pediatric surgery service, which is responsible for coordinating care among managing specialty services (eg, neurosurgery or orthopedic surgery).

The Johns Hopkins Medicine Institutional Review Board approved the study (IRB00076900).

Data Set

Data were extracted from the pediatric trauma registry and the EHR data warehouse (ie, the Clarity database of Epic). We limited EHR data to encounters with trauma activation levels of alpha, bravo, and critical trauma transfers that were managed between January 1 and December 31, 2017. Demographic and encounter data including age, sex, origin of patient, trauma activation level, injury severity score, and Glasgow Coma Scale score were collected from the registry. Admission, discharge, and transfer (ADT) data and metadata of 5 clinical activities (ie, notes, procedure orders, medication orders, flow sheet entries, and medication administration entries) captured in the EHR were collected from the EHR data warehouse. For each EHR activity type, we obtained the encounter ID (visit ID), activity timestamp, unique ID, and generic clinical roles (eg, attending or resident) of the HCP that performed the activity.

The note metadata included the service of the authors, whereas the procedure orders, medication orders, and medication administration entries included the care location (eg, ED or PICU) where the activity was performed.

Data Preparation

Each encounter was assigned a randomly generated, unique study ID. Timestamps of EHR metadata were normalized by replacing them with time (in minutes) from ED arrival, which ensured that the temporal sequence of events was maintained for each encounter. Activities without a full complement of data were excluded. Activities that were initiated by the EHR system and initiated by student roles (eg, nursing and medical students) were also excluded as they bore no accountability for patient care. As notes were typically signed off much later from when they were started, we considered the note creation time as the note completion time. As flow sheet and note data lacked care location data, we inferred the care location for each activity from the ADT data as follows: First, a location timeline was generated from the ADT data (ie, sequence of admissions to various hospital locations from ED arrival to hospital discharge). The normalized timestamps of each activity in the flow sheet and note metadata were then subsequently related to the location timeline, and the corresponding care location was taken as the care location where the flow sheet and note activities were performed.

Identification of Functional Roles

We considered collaboration at the level of functional roles (eg, ED nurse, neurosurgery resident, PICU fellow, and surgery attending) rather than individuals, as past studies have shown that mirrors the reality of clinical practice [25]. To determine functional roles, we identified the service (eg, orthopedic or ophthalmology service) to which each identified HCP belonged and prefixed it to their generic role (eg, resident or attending). This service could be a service that is bound to a care location (eg, ED, PICU, or general care floor) or a service that operates across care locations (eg, general pediatric surgery service or physical therapy).

We assumed that the services of certain functional roles (eg, attending, fellows, physician assistants, and nurse practitioners working on specialty services) were fixed as determined from their notes. Chart reviews and directory lookups were conducted to identify the services of individuals whose services could not be determined from the extracted metadata. The services of medical residents, which frequently change as they rotate through various services for their training, were determined on an encounter basis derived from the service of the attending that cosigned the notes. The services of registered nurses, unit-based nurse practitioners, and allied HCPs (excluding radiology technicians) were determined by taking the mode of the frequency distribution of the location of the activities they performed. The services of radiology technicians were determined on an encounter basis similar to that of residents.

Activities by individuals whose services could not be determined were excluded. Since the location of flow sheet and note activities were inferred, the records were excluded if the inferred location did not correspond to the base unit of HCPs.

Methodologic Approach

We used a process mining approach, which is a field of data science that *aims to discover, monitor, and improve real processes by extracting knowledge from event logs* [26]. The starting point for process mining is an event log, which contains a collection of events. Each event represents a discrete activity (eg, note writing) in a given process (eg, clinical care), performed by an actor (eg, ED resident), and relates to a case (eg, patient encounter). Each event is time-stamped (eg, order placed on January 22, 2000, at 10:45 AM), allowing all events for a patient encounter to be ordered chronologically [27]. By applying a *metric* described below, social network interactions and collaboration between different functional roles were obtained [28].

Working together is a commonly used metric for representing collaboration in unstructured processes with frequent ad hoc behavior such as in health care [29]. The working together metric counts how frequently 2 actors work together on same cases [28]. In its regular form, the working together metric does not accommodate for temporal distance between actors, which is important in health care where different HCPs are involved in patient care at different stages of care. Consequently, we defined a variant of the working together metric, referred to as *working closely together*, to account for temporal distance among actors. The working closely together metric counts the number of times 2 actors worked closely together with respect to time for a given patient relative to the number of times the 2 actors had the opportunity to work together. To operationalize this metric, we considered the shift rotation as the unit of clinical work and collaboration, and assumed that functional roles that were involved in the care of a patient during a shift had the opportunity to work together, whereas functional roles that were captured in the EHR within a similar time interval were *working closely together*. Therefore, this metric translates to functional roles that are jointly involved in completing the same tasks or completing disparate tasks within the same time interval.

Generation of Master Event Log

EHR metadata were processed into an event log consisting of the study ID, normalized time, EHR activity type, unique ID, and functional role of the HCP, and care location. Multiple *same-time events* were generated from notes, procedures, and medication orders that involved multiple HCPs. The encounter timeline was divided into shift rotations (day: 7 AM-6:59 PM and night: 7 PM-6:59 AM) numbered 0 to N, and each event in the event log was labeled with the corresponding shift number and shift type (day or night). Events within each shift were partitioned into segments based on *natural breaks* in the continuity of events. We assumed a natural break to be a minimum of 30 minutes between adjacent events in the event log to accommodate the lag between the occurrence of activities in real life and registration in the EHR. The Jenks Natural Break Optimization algorithm [30] was used to determine the optimal break interval between 30 and 120 minutes in 5-minute increments.

Generation of Sublogs

The master event log was divided based on shift type (day or night) and care location (ED, floor, or PICU) to obtain six individual event logs: ED morning, ED night, floor morning, floor night, PICU morning, and PICU night.

Network Representation

For each individual sublog, an undirected edge (ie, the relationship among nodes) was created for all pairwise combinations of identified functional roles within each event segment. Unique edges across all segments across all shifts across all encounters were obtained as the collaboration network. The weight of the edges was obtained by dividing the number of shifts an edge was present between 2 functional roles by the number of shifts in which both functional roles were involved, which effectively normalized the weights and accommodated for variation in care team composition across encounters.

Threshold Selection

To prevent the capture of spurious edges (ie, edges that do not really exist or edges with spurious weights) in network analysis, a threshold number of shared encounters among nodes (ie, functional roles) is usually applied to constructed networks. The eventual network structure is sensitive to the selected threshold. Various approaches that have been used to determine this threshold are subjective [31], including arbitrary selection [32], clinician informed [33], and retaining only a fixed top percentage of the strongest edges [34]. In this study, we attempted to take a more objective approach to threshold determination by introducing a heuristic method akin to the elbow method [35], which is used to determine the optimal number of clusters in k-means clustering. For each event log, we obtained and plotted the rate of change of the total number of edges removed as the threshold value (ie, representing the number of shared shifts) was incrementally increased from 2 to 20 and obtained a LOWESS (Locally Weighted Scatterplot Smoothing)-smoothed curve of the plot. The elbow point—the smallest threshold value at which the rate of change becomes insignificant or constant, was taken as the optimal threshold. The underlying assumption of this method is that as the threshold of the shared number of encounters is increased, trivial and spurious edges are removed, and the network structure changes up to a point where further increases in threshold value result in minimal removal of edges with little or no change in the network structure. At this threshold point, we assume that the network structure is relatively stable and only significant edges and nodes remain.

Network Visualization and Analysis

We used the igraph 1.1.1 package [36] in R (version 3.4.0; R Foundation for Statistical Computing) [37] to create and visualize the networks. From each network, we obtained the node count (ie, number of functional roles) and edge count (ie, number of relationships among functional roles). We used the linkcomm package 1.0.11 [38] to identify the overlapping communities in the networks. A community is a subnetwork that contains a high density of edges among members but fewer edges with members of the larger network, thus represents a tightly knit subgroup [39]. The linkcomm package is an R implementation of the algorithm by Ahn et al [40] that, as opposed to other community detection algorithms that cluster nodes—clusters edges assuming a node can belong to multiple communities, thus enabling the discovery of overlapping and nested communities. The algorithm by Ahn et al [40] is the most commonly used overlapping community detection algorithm and tends to produce superior performance if multiple ad hoc behaviors result in a high degree of overlap in derived networks, as is commonly seen in health care settings [41,42]. The algorithm uses a hierarchical clustering method to produce a dendrogram that, in the default setting, is cut at a level that maximizes the partition density [40]. The linkcomm package offers a unique visualization that uses different colors to depict edges and nodes that belong to different communities. Nodes are sized to reflect the number of communities the node belongs to, with larger nodes belonging to more communities. Nodes belonging to more than one community are also presented as pies with the pies divided and colored based on the proportion of the edges for that node in various communities that the node belongs. We parameterized the algorithm with the McQuitty hierarchical clustering method, also known as the Weighted Pair Group Method with Arithmetic Mean [43], so that edge weights can be considered in community determination. We subsequently obtained community-depicted networks produced at maximum modularity that were visualized with easily understandable network layout algorithms.

Statistical Analysis

We obtained and compared descriptive statistics of demographic, injury, and outcome characteristics of day and night shift encounters. We also compared composition of days and night shift event logs for each care location. Differences among interval and categorical variables were examined using Wilcoxon rank-sum and Pearson chi-square tests, respectively. Differences were considered statistically significant at an $\alpha<.05$. The analysis was performed using Stata 13 [44].

Validation

Two forms of validation were conducted. In the first validation step, we compared the results of this study with the secondary analysis of data from and results of a previous study [45] in which we developed a *role-location matrix*, which is a 2×2 table of functional roles and the inpatient locations in which they typically worked via semistructured interviews with clinicians (n=21) and subject matter experts (n=22), and a review of the institutional and trauma registry protocol. We compared the functional roles and the locations in which the functional roles were found in this study to the role-location matrix. In the second validation step, we validated the collaboration patterns of pediatric trauma MTS via member-checking interviews (n=6) with care professionals (ie, pediatric trauma program director, PICU attending, and pediatric trauma nurses) that were involved in pediatric trauma care. The interviews were conducted by AD, KW, and GSD and APG as a group. During each session, the collaboration patterns of care teams were individually presented to the HCP, who were asked to comment on (1) the accuracy and completeness of the roles that were captured by location and shift; (2) whether the collaborative patterns mirrored reality or not; and (3) whether the differences between day and night

patterns for a given care location (ie, ED, PICU, or floor) were suggestive of reality.

Results

Overview

There were 413 encounters in the cohort, of which 65.9% (272/413) and 34.1% (141/413) began during day and night shifts, respectively. Compared with patients who arrived during day shifts, those who arrived during night shifts were significantly older (median age 7 vs 10 years; *P*=.04), had a higher proportion of critical trauma transfers (8.8% vs 26.2%; *P*<.001), and had a higher proportion of penetrating injuries (5/272, 1.8% vs 11/141, 7.8%; *P*<.001; Table 1). There were no significant differences in sex, injury severity score, Glasgow Coma Scale, operating room and PICU admissions, ED, PICU, hospital length of stay, and mortality.

Table 1. Comparison of demographic and encounter characteristics by shift type[a].

Variables	Day (n=272)	Night (n=141)	*P* value
Age (years), median (IQR)	7 (3-11)	10 (3-13)	.04
Male sex, n (%)	184 (67.7)	83 (58.9)	.08
Trauma activation, n (%)			<.001
Alpha	26 (9.6)	5 (3.6)	
Bravo	222 (81.6)	99 (70.2)	
Critical trauma transfer	24 (8.8)	37 (26.2)	
Origin, n (%)			<.001
Scene of injury	245 (90.1)	102 (72.3)	
Transfer	2 (0.7)	38 (27)	
Others	2 (0.7)	1 (0.7)	
Injury type, n (%)			.01
Blunt	259 (95.2)	126 (89.4)	
Penetrating	5 (1.8)	11 (7.8)	
Others	8 (2.9)	4 (2.8)	
ISS[b], median (IQR)	5 (2-10)	5 (2-9)	.76
GCS[c], median (IQR)	15 (15-15)	15 (15-15)	.48
ED[d] LOS[e] (minutes), median (IQR)	253.5 (187-361)	254 (146-374)	.52
OR[f] admission, n (%)	41 (15.1)	22 (15.6)	.89
PICU[g] admission, n (%)	43 (15.8)	27 (19.2)	.39
PICU LOS (days), median (IQR)	1 (1-3)	1 (1-2)	.48
Hospital LOS (hours), median (IQR)	7 (4-32)	14 (4-41)	.21
Mortality, n (%)	7 (2.6)	2 (1.4)	.72

[a]Day shift is defined as 7 AM to 6:59 PM, whereas night shift is defined as 7 PM to 6:59 AM.

[b]ISS: injury severity score.

[c]GCS: Glasgow Coma Scale.

[d]ED: emergency department.

[e]LOS: length of stay.

[f]OR: operating room.

[g]PICU: pediatric intensive care unit.

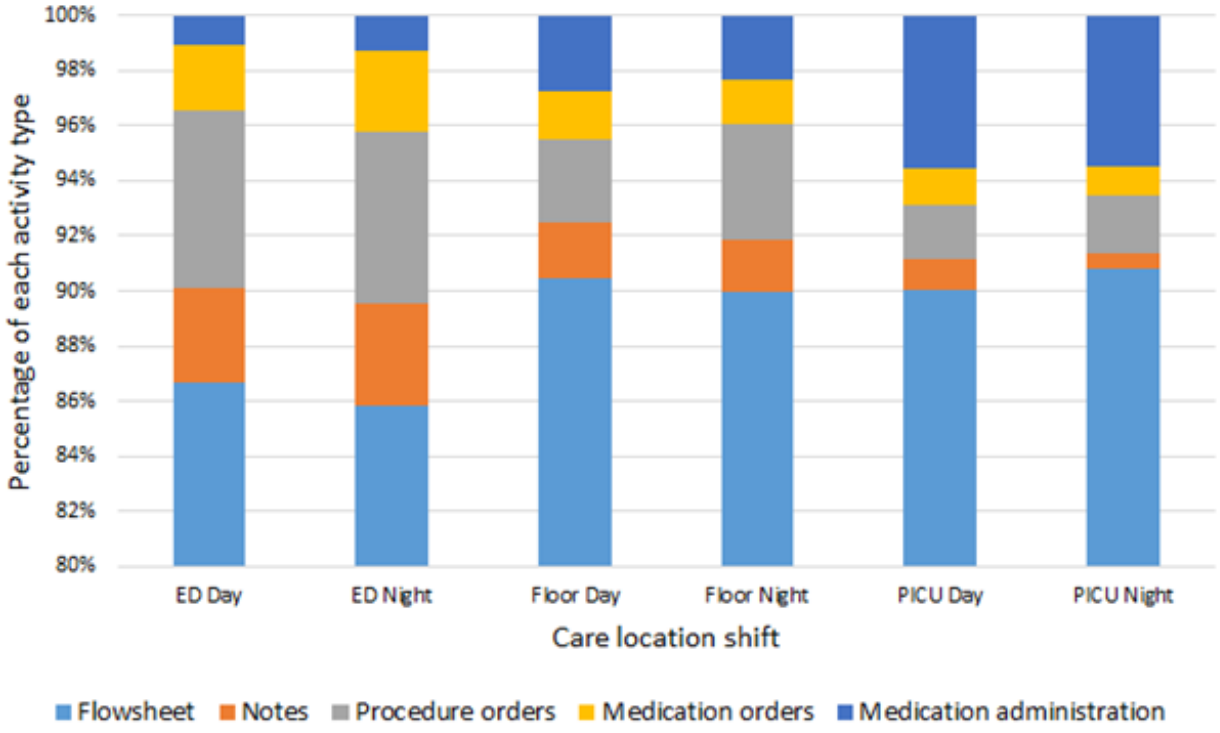
Master Event Log Characteristics

There were 837,318 events in the initial event log, respectively. Only 0.19% (1564/837,318) of the events were excluded owing to the inability to resolve the functional role of the actor. Consequently, 835,754 events remained in the master event log. Flow sheet entries accounted for 89.45% (749,000/837,318) of all events in the log. A total of 1647 unique HCPs occupying 110 functional roles were identified, of which 58 functional roles were recorded in at least 4.8% (20/413) of encounters. The ED registered nurses were recorded in all 413 encounters, whereas the ED attending, ED resident, and ED radiology technician were recorded in 98.5% (407/413), 93.2% (385/413), and 80.6% (333/413) encounters, respectively.

Comparison of Sublogs Obtained Based on Shift Type and Care Location

Figure 1 depicts the composition of the individual sublogs for each care location and shift duty. The proportions of various activities in the day and night logs for each care location were similar, with some notable differences. The ED night log contained more medication administration orders than the ED day log, which contained more flow sheet events. The floor day log contained more medication administration than the floor night, which contained more procedure-order events. The PICU day contained more notes events that the PICU night, which contained more flow sheet events.

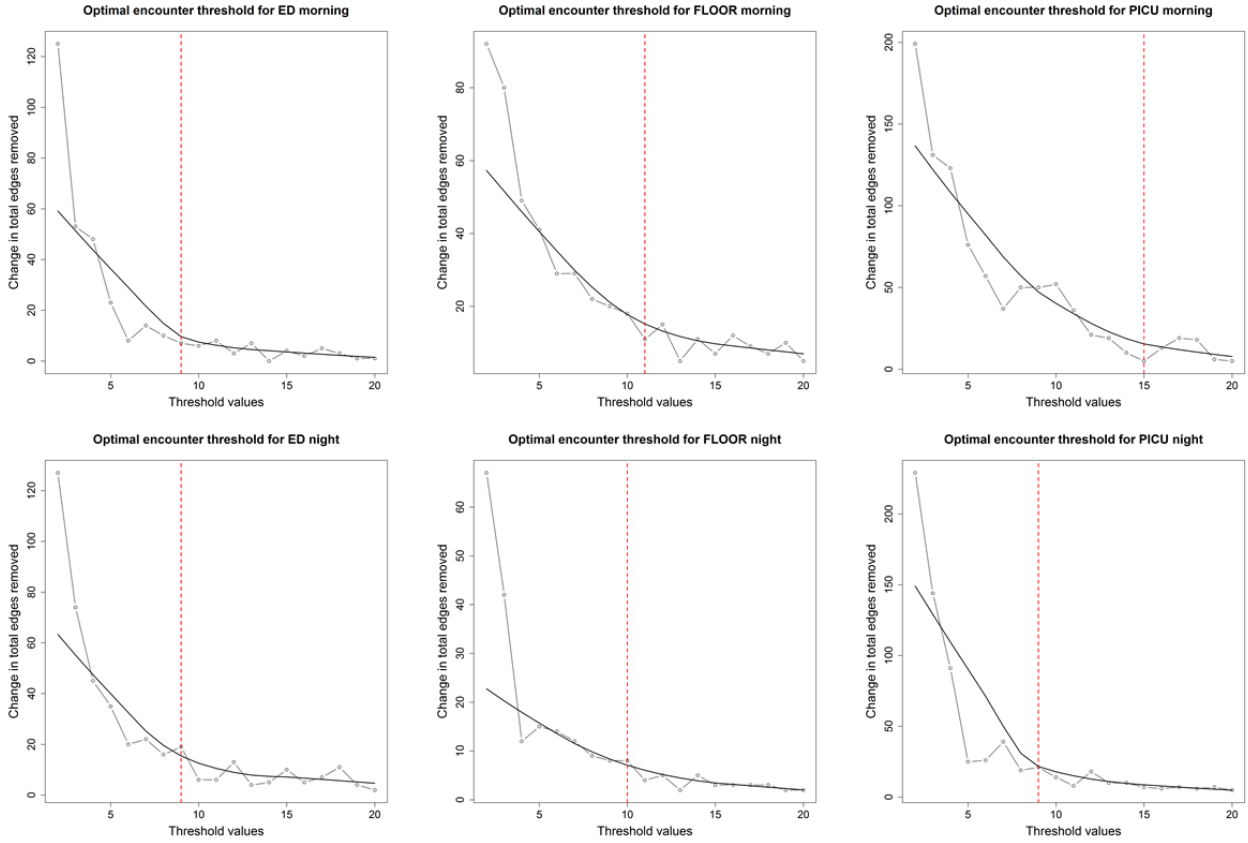
Figure 1. Comparison of the composition of various activity types by care location and shift type. ED: emergency department; PICU: pediatric intensive care unit.



Threshold Selection

Figure 2 shows the plots of the rate of change of total edges removed against increasing threshold values. The gray line and point plot show the difference in edges removed as the threshold is increased, whereas the smooth black line is the LOWESS curve. Some LOWESS curves, such as the ED morning and PICU night, have sharply defined elbows, whereas others have subtle elbows. The red vertical lines indicate the selected threshold number of shared encounters by HCPs for each event log. For both ED day and night, the threshold was determined to be 9. For the floor, 11 and 10 were selected as the thresholds for day and night, respectively, whereas for the PICU, 15 and 9 were selected as the day and night thresholds, respectively.

Figure 2. Determination of encounter threshold for each event log. ED: emergency department; PICU: pediatric intensive care unit.



Collaborative Care Teams in the Pediatric ED

Figure 3 shows the collaborative care team pattern in the ED during the day and at night visualized using the Kamada Kawai layout algorithm [46], which is a force-directed algorithm. Table 2 contains the meaning of the abbreviations used in Figure 3 and in all other network diagrams in this paper. The day pattern contained 18 nodes and 87 edges, whereas the night pattern contained 28 nodes and 160 edges. The night pattern was distinctively star-shaped and had 5 overlapping communities with the ED attending, residents, nurses, radiology technicians, and the general pediatric surgery attending and resident forming the core and belonging to all 5 communities. The day pattern had a less distinctively defined star pattern and had only 1 community. Attending-resident pairs from neurosurgery and orthopedic surgery services, and allied HCPs, including social workers, chaplains, and child life specialists, were at the periphery in both patterns. Attending-resident pairs from otolaryngology and plastic surgery were seen only in the night pattern and belonged to separate communities, whereas only the resident from the ophthalmology service was seen in the night pattern. The PICU nurse, resident, and the imaging data coordinator (IDC) were also seen in the night pattern.

Figure 3. Collaborative care team patterns in the emergency department. Left: day shift; right: night shift.

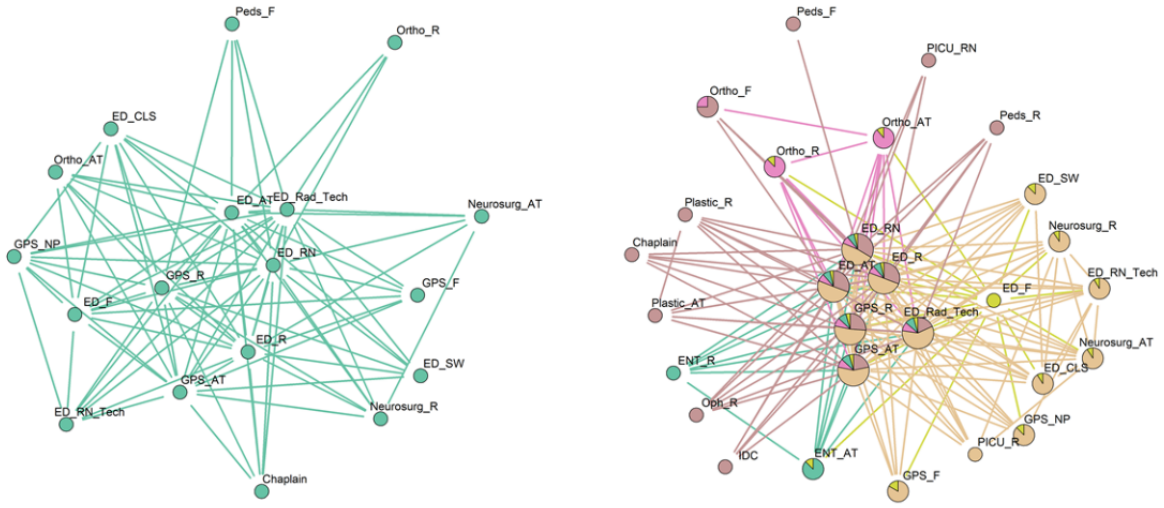


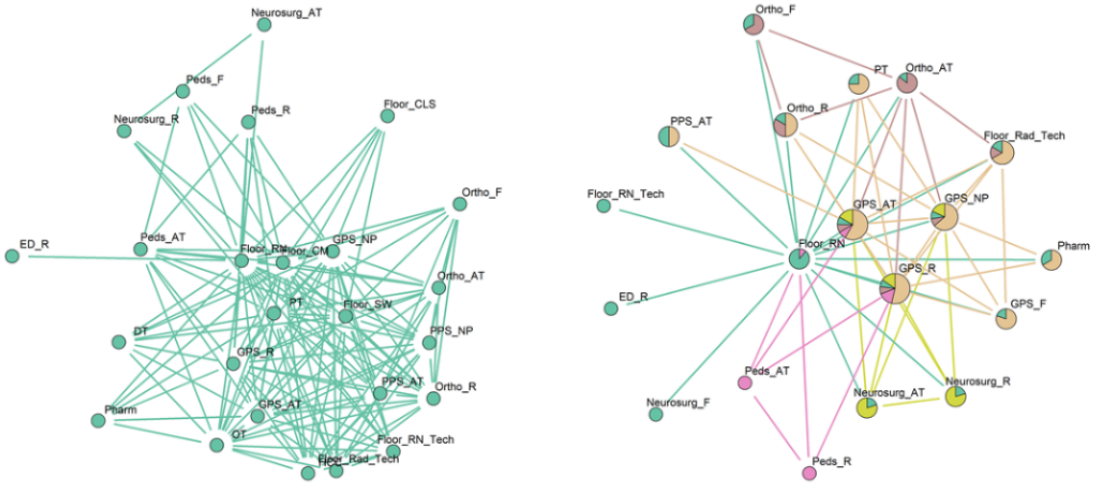
Table 2. Abbreviations used in the network diagrams.

Abbreviation	Meaning
Anes	Anesthesia
AT	Attending
CLS	Child life specialist
CM	Case manager
DT	Dietitian
ED	Emergency department
F	Fellow
GPS	General pediatric surgery
HCC	Home care coordinator
IDC	Imaging data coordinator
Neuro	Neurology
Neurosurg	Neurosurgery
NP	Nurse Practitioner
Oph	Ophthalmology
Ortho	Orthopedic surgery
OT	Occupational therapy
PA	Physician assistant
Peds	Pediatrics
Pharm	Pharmacist
PICU	Pediatric intensive care unit
PMR	Physical medicine and rehabilitation
PPS	Pediatric pain service
PT	Physical therapist
R	Resident
Rad_Tech	Radiology technician
RN	Registered nurse
RN_Tech	Nurse technician
SW	Social work

Collaboration Patterns of Care Teams in the Floor

Figure 4 shows the collaboration pattern on the floor during the day and at night visualized using the large graph layout [47]. The day pattern contained 24 nodes and 135 edges, whereas the night pattern contained 19 nodes and 55 edges. The bedside nurse was at the center of both patterns. Functional roles present in the day pattern but absent in the night pattern were home care coordinators, case managers, social workers, child life specialists, occupational therapy, and dietitians. The ED resident was present in the night pattern but not in the day pattern. One community was identified in the day pattern, whereas 5 overlapping communities were identified in the night pattern with the neurosurgery, orthopedic surgery, and pediatric services having separate communities and the general pediatric surgery-resident and general pediatric surgery attending belonging to all 5 communities.

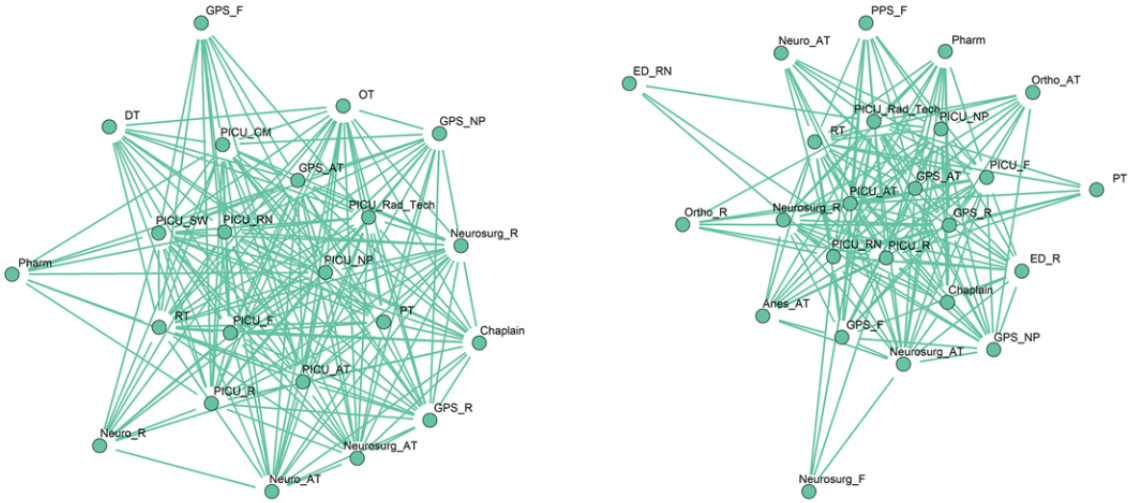
Figure 4. Collaborative care team pattern on the floor. Left: day; right: night. Only one community was identified in the day pattern while 5 communities (different colors) were identified in the night pattern.



Collaboration Patterns of Care Teams in the PICU

Figure 5 shows the day and night collaborative care team pattern in the PICU visualized using the Frutchterman–Reingold layout algorithm [48]. The day pattern contained 30 nodes and 283 edges, whereas the night pattern contained 24 nodes and 175 edges. Both collaboration patterns had a large spherical core made up of functional roles from the PICU, general pediatric surgery, neurosurgery, and neurology services (day pattern only), and few *appendages* that include functional roles from the orthopedic surgery, ophthalmology and pediatric pain service. One community was identified in both the patterns. Functional roles present in the day pattern but absent in the night pattern were unit case managers, social workers, occupational therapists, and dietitians. Functional roles present in the night pattern but absent in the day pattern include the ED resident, ED nurse, orthopedic surgery team, and anesthesiology attending.

Figure 5. Collaborative care team patterns in the pediatric intensive care unit. Left: day; right: night.



Validation

In a previous study that used semistructured interviews with care professionals [45], we identified 56 roles involved in pediatric trauma care across all care locations. In this study, we identified a total of 110 functional roles and 58 frequent functional roles across all locations. Eight functional roles were identified in a previous study but not in this study. These roles were ED documenting nurses, charge nurses, emergency medical services personnel, security, family or caregiver, pediatric trauma manager, perfusionist, and in-hospital transport team. A total of 54 functional roles were identified in this study but not in a previous study. Most of these roles belonged to specialty service roles that were not frequently involved in patient care. Of the 58 frequent roles identified in this study, 15 (26%) were not identified in the prior study. These roles included dietitians, IDCs, home care coordinators, ophthalmology service, otolaryngology service, neurology service, pediatric pain service, and plastic surgery services.

A comparison between this study and the prior study showed that the locations of the functional roles they had in common mostly matched. For example, both studies confirmed that ED nurses, ED residents, and PICU nurses go to the floor, usually during patient transport. However, few differences exist. For example, in a previous study, it was revealed that the PICU attending, PICU fellow, and respiratory therapist responded to

alpha traumas in the ED both during the day and at night, but this was not captured in this study.

The 6 HCPs who were interviewed for this study found the composition of the derived care teams to be largely accurate. However, they pointed out that some functional roles were not accurately captured. For example, PICU attendance, PICU fellows, and respiratory therapists were not identified in the ED collaborative care teams (Figure 3). This was attributed to the fact that PICU team members who responded to traumas rarely did any documentation while in the ED. They also pointed out that the team pattern for the PICU night did not capture the ED social worker who usually covers the PICU at night, and the floor care team patterns were missing functional roles from anesthesiology.

Regarding interactions among roles and communities that were identified, clinicians confirmed the general pediatric surgery coordinated care among specialist services and understood why they belonged to multiple communities. Clinicians explained why only 1 community was identified in the PICU, as concerted efforts have been made to improve coordination of care between the PICU and surgical services, and the PICU characteristically performed multidisciplinary rounds with other nonsurgical services and allied HCPs. However, clinicians acknowledged that collaboration with the orthopedic service, particularly in the PICU, can be further improved. Clinicians confirmed that the neurosurgery service was well integrated into the trauma team in the ED.

Discussion

Principal Findings

We compared diurnal differences in the composition and organization of collaborative care teams at 3 care locations in a level I pediatric trauma center using EHR data. Our study is unique in several ways. First, we introduced a heuristic for determining the threshold number of shared patient encounters for interaction between HCPs. The heuristic method allows a more objective approach to threshold selection. In 67% (4/6) of the scenarios, we obtained distinct elbow points, whereas in the other 33% (2/6) of the scenarios, we easily identified the appropriate threshold on closer examination. Second, we used an overlapping community detection algorithm that allowed a functional role to be part of multiple communities to reflect ad hoc clinical collaborations that clinicians form to address the unique needs of patients. In 33% (2/6; ED night and floor night) of the scenarios, we identified multiple overlapping communities suggestive of MTS, whereas in the other 67% (4/6), only a single community was identified. Third, we confirmed the presence of MTS using the EHR data. We also showed that the EHR data complemented interview data for identifying functional roles. Although interview data were especially helpful in identifying team members or roles who rarely document in the EHR, the EHR data enabled a more comprehensive and systematic analysis and identification of functional roles (56 functional roles identified with interviews vs 110 with EHR data).

There were 3 significant differences among patients who arrived during night shift compared with those who arrived during the day time. The patients who arrive at night tend to be older (median age 10 vs 7 years) and have penetrating injuries (11/141, 7.8% vs 5/272, 1.8%). This is likely related to prevailing epidemiological conditions and is consistent with what has been reported in the literature [49,50]. These patients also tended to arrive as transfers from other facilities (38/141, 27% vs 2/272, 0.7%). A higher percentage of transfers received at night reflects operational circumstances. Our pediatric trauma center is a level I trauma center that serves as a referral center for a large area. The decision to transfer patients is made by the originating facility, but several factors determine when patients physically arrive at our facility. First, the patients must be stabilized (to some extent) at the originating facility to ensure that they will survive transportation before departing the originating facility. Second, the level of staffing at the originating facility may influence transfer decisions such that patients who would be unsafe to manage at night when staffing is low are transferred to us after stabilization. Third, the distance of the originating facility and logistics of transportation can influence when transfer patients physically arrive at our facility.

There were some notable differences in the composition of event logs at various locations. The lower proportion of flow sheet activities in the ED is likely because of the relatively short time (usually <60 minutes) spent in the ED as compared with the entire hospital stay (usually days). The higher proportion of flow sheets and medication activities in the PICU compared with the ED and the floor reflects the intensive care provided in the PICU. The higher proportion of procedure orders in the ED compared with both the floor and PICU suggests the initiation and delivery of immediately necessary and likely lifesaving interventions.

Important differences were observed between the collaborative care teams in the ED during the day and at night. Compared with the day pattern, the night pattern had a better-defined core team made up of ED and general pediatric surgery personnel and involved more specialty services, which was reflective of the nature and severity of injury of patients presenting at night [8,51,52]. In addition, the neurosurgery team was part of both day and night patterns. However, in the night pattern, the neurosurgery team was part of the main community that included the core team and allied HCPs. This suggests that the neurosurgery team has a close relationship with the trauma team in the ED, which was confirmed by the interviewed clinicians. In addition, the collaborative care team in the ED at night included roles that did not exist in the day pattern. These roles include the IDC, a role that is responsible for uploading imaging data from transferring hospitals that do not use an interoperable EHR (which can be explained by the significantly higher number of trauma transfers arriving at night), and the PICU resident and PICU nurse, which suggested greater involvement at night, possibly to facilitate faster admission to the PICU). The orthopedic surgery attending and resident, and the ED resident and ED nurse were captured by the night pattern in the PICU but not during the day, which suggested greater involvement in PICU-related activities of trauma patients at night.

Compared with the day pattern, multi-team structures were more pronounced at night. Constituent specialty teams usually consisted of attending-resident pairs, except for the

ophthalmology service, which consisted only of residents. Validation with clinicians confirmed that ophthalmology attending physicians do not take in-house night duty calls, given the seldom emergent nature of many ophthalmologic problems. Conspicuous multi-team structures reflected the presence of fewer ancillary support services that often serve as coordinators of care. In the ED, as ancillary support services were present at night, this may be reflective of the greater need of the patients received at night and the difficulty in coordination of care among the various services. On the floor, where ancillary support services are not present at night, this suggests that ancillary support staff play important roles in coordinating care and ensuring that various teams function as a unit. However, this was not the case in the PICU, where the night collaboration pattern was essentially similar to the day pattern despite the absence of ancillary support staff at night.

There are a number of reasons for the observed variations between day and night networks. As described, the level of staffing during the day was higher than that at night. During the day, more functional roles and support staff (care coordination, social work, etc) are present, and they participate in collaboration between teams. At night, some functional roles and nonessential staff are not available, which changes the dynamics of work and collaboration. In addition, more care activities (patient rounds, elective procedures, discharge planning, etc) occur during the day as opposed to nighttime. These activities create the need and opportunity for close collaboration compared with nighttime. Finally, there are likely differences in the manner of collaboration, for example, the use of non-EHR–based communications such as telephone and paging is more common during the night when team members tend to be more geographically dispersed, as opposed to during the day when they are geographically closer or physically working together.

In addition to organizational factors, methodological issues may also account for the variations. Only a single community was identified at all locations during the day. Although it is probable that the specialty teams actually do work very closely together during the day, it is likely that they do in a multi-team setup, which we did not identify by overlapping community detection. This may be owing to several reasons. The data may lack adequate *power* to detect overlapping communities during daytime. Certainly, smaller teams could be identified using cliques, which are unique subnetworks containing at least three nodes that are all connected to one another by edges. However, given the number of nodes involved, hundreds of overlapping cliques would be identified, which would be difficult to interpret. Another factor could be threshold selection. It is possible that by selecting different thresholds for day networks, we could identify overlapping communities. However, because the weights of the edges are considered by the community detection algorithm, such a sensitivity analysis was not required. Nevertheless, a narrow sensitivity analysis of the day networks using +/−1 the selected threshold did not show any major difference in terms of communities (Multimedia Appendices 1-3).

This study has several implications: the methodology can be adapted and used in other settings to identify and study MTS structures in an efficient manner. The methodology can also be adapted to study how MTS evolves over the care timeline of patients and identify areas in need of improvement. In-depth analysis of MTS across time, location, and team members using EHR metadata can provide insights to support management and operational decisions. For example, it can be used to derive insights into how HCPs and care teams organize themselves given the realities of *actual work*, rather than how they are supposed to organize according to protocols. Such insights can be used to inform staffing and team composition decisions, team training and development efforts, and complement efforts to improve collaboration and coordination to improve team-based health care delivery. In addition, the ability to compare temporal patterns in MTS dynamics based on EHR metadata enables assessment and evaluation of the impact of any quality improvement and intervention efforts aimed at improving MTS performance.

This study had several notable limitations. First, by only EHR data, we did not capture other important teamwork-related activities such as face-to-face and telephone conversations, which are a major part of clinical activities [53]. Second, we were less likely to capture functional roles that documented infrequently in the EHR. For example, we were unable to capture the PICU attending and PICU fellows in the ED patterns for both day and night, as these 2 roles rarely used the EHR for documentation while in the ED. We were also unable to capture several other HCPs, such as emergency medical services personnel, security, family and/or caregiver, pediatric trauma manager, perfusionist, and in-hospital transport team who rarely or never use the EHR for a trauma case, but are an integral part of the care team, as revealed by interview data. Other methods, such as in-depth interviews or direct observations, can be used to overcome these limitations. Third, our method for determining the functional roles was based on heuristics. Consequently, it is possible that not all possible roles were identified, and that some of the assigned functional roles were inaccurate. Nevertheless, as demonstrated, the methodology performs quite well; future EHR systems should be designed to support functional roles, which are the appropriate unit of clinical collaboration, rather than individuals; for example, clinical documentation could be primarily based on functional roles, but signed as individuals. Such systems have the potential to optimize collaborative work to deliver improved care and enable robust research using EHR data.

Conclusions

We identified and described diurnal variations in MTS and collaborative care teams at various locations and stages of care, as well as various shift types in a pediatric trauma center using EHR data. We validated our results using qualitative data and showed that the derived structures can accurately represent reality. The methodology described can be adapted to study how MTSs evolve over time and across locations, and the insights can be used to support management and operational decisions.

Acknowledgments

This study was supported by the Agency for Healthcare Research and Quality (R01HS023837; principal investigator: AG). The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality. The authors thank Cathi Habertson and Laura Fortmann for their role in this research.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sensitivity analysis of the collaborative care team patterns for the emergency department day shift using +/−1 the selected threshold. At a lower threshold of 8, a minor overlapping community that included the orthopedic resident (Ortho_R) and imaging data coordinator (IDC) as additional members was identified.
[PNG File , 772 KB - jmir_v24i2e30351_app1.png]

Multimedia Appendix 2
Sensitivity analysis of the collaborative care team patterns for the floor day shift using +/−1 the selected threshold. At a threshold of 10, functional roles with single connection to the floor nurse were identified. However, these functional roles were removed as the threshold is increased with no functional role having a single connection at a threshold of 12.
[PNG File , 712 KB - jmir_v24i2e30351_app2.png]

Multimedia Appendix 3
Sensitivity analysis of the collaborative care team patterns for the pediatric intensive care unit day shift using +/−1 the selected threshold. At a higher threshold of 16, a minor community including the neurology resident as the only additional member was identified.
[PNG File , 855 KB - jmir_v24i2e30351_app3.png]

References

1. DiazGranados D. Understanding patient care as a multiteam system. In: Pushing the Boundaries: Multiteam Systems in Research and Practice. United Kingdom: Emerald Group Publishing Limited; 2014.
2. TeamSTEPPS 2.0. Agency for Healthcare Research and Quality. URL: http://www.ahrq.gov/teamstepps/instructor/index.html [accessed 2022-01-27]
3. Mathieu J, Marks M, Zaccaro S. Handbook of Industrial, Work & Organizational Psychology - Volume 2: Organizational Psychology. Newbury Park, California: SAGE Publishing; 2001.
4. Multiteam systems. In: Handbook of Industrial, Work & Organizational Psychology - Volume 2: Organizational Psychology. Newbury Park, California: SAGE Publications; 2001.
5. Shuffler ML, Carter DR. Teamwork situated in multiteam systems: key lessons learned and future opportunities. Am Psychol 2018;73(4):390-406. [doi: 10.1037/amp0000322] [Medline: 29792456]
6. Resources for optimal care of the injured patient 2014/resources repository. American College of Surgeons. 2014. URL: https://www.facs.org/quality-programs/trauma/tqp/center-programs/vrc/resources [accessed 2022-01-27]
7. Tannenbaum SI, Mathieu JE, Salas E, Cohen D. Teams are changing: are research and practice evolving fast enough? Ind Organ Psychol 2015 Jan 07;5(1):2-24. [doi: 10.1111/j.1754-9434.2011.01396.x]
8. Vaziri K, Roland J, Robinson L, Fakhry S. Optimizing physician staffing and resource allocation: sine-wave variation in hourly trauma admission volume. J Trauma 2007 Mar;62(3):610-614. [doi: 10.1097/TA.0b013e31803245c7] [Medline: 17414336]
9. Parsch W, Loibl M, Schmucker U, Hilber F, Nerlich M, Ernstberger A. Trauma care inside and outside business hours: comparison of process quality and outcome indicators in a German level-1 trauma center. Scand J Trauma Resusc Emerg Med 2014 Oct 31;22:62 [FREE Full text] [doi: 10.1186/s13049-014-0062-2] [Medline: 25366718]
10. Di Bartolomeo S. The 'off-hour' effect in trauma care: a possible quality indicator with appealing characteristics. Scand J Trauma Resusc Emerg Med 2011 Jun 09;19:33 [FREE Full text] [doi: 10.1186/1757-7241-19-33] [Medline: 21658243]
11. Laupland KB, Ball CG, Kirkpatrick AW. Hospital mortality among major trauma victims admitted on weekends and evenings: a cohort study. J Trauma Manag Outcomes 2009 Jul 27;3:8 [FREE Full text] [doi: 10.1186/1752-2897-3-8] [Medline: 19635157]
12. Hollnagel E. Why is work-as-imagined different from work-as-done? Resilient Health Care 2017;2:279-294. [doi: 10.1201/9781315605739-24]
13. Crawford ER, LePine JA. A configural theory of team processes: accounting for the structure of taskwork and teamwork. Acad Manag Rev 2013 Jan;38(1):32-48. [doi: 10.5465/amr.2011.0206]

14. Innovation in scientific multiteam systems: confluent and countervailing forces. George Tech. URL: https://sites.nationalacademies.org/cs/groups/dbassesite/documents/webpage/dbasse_084357.pdf [accessed 2022-01-28]
15. Burke C. A look into the challenging world of MTS data collection. In: Pushing the Boundaries: Multiteam Systems in Research and Practice. United Kingdom: Emerald Group Publishing Limited; 2014.
16. Luciano MM, Mathieu JE, Park S, Tannenbaum SI. A fitting approach to construct and measurement alignment. Organizational Res Methods 2017 Sep 13;21(3):592-632. [doi: 10.1177/1094428117728372]
17. Chen Y, Lorenzi NM, Sandberg WS, Wolgast K, Malin BA. Identifying collaborative care teams through electronic medical record utilization patterns. J Am Med Inform Assoc 2017 Apr 01;24(e1):e111-e120 [FREE Full text] [doi: 10.1093/jamia/ocw124] [Medline: 27570217]
18. Gray JE, Feldman H, Reti S, Markson L, Lu X, Davis RB, et al. Using Digital Crumbs from an Electronic Health Record to identify, study and improve health care teams. AMIA Annu Symp Proc 2011;2011:491-500 [FREE Full text] [Medline: 22195103]
19. Wooldridge AR, Carayon P, Hoonakker P, Hose B, Ross J, Kohler JE, et al. Complexity of the pediatric trauma care process: implications for multi-level awareness. Cogn Technol Work 2019 Aug;21(3):397-416 [FREE Full text] [doi: 10.1007/s10111-018-0520-0] [Medline: 31485191]
20. Wooldridge AR, Carayon P, Hoonakker P, Hose B, Eithun B, Brazelton T, et al. Work system barriers and facilitators in inpatient care transitions of pediatric trauma patients. Appl Ergon 2020 May;85:103059 [FREE Full text] [doi: 10.1016/j.apergo.2020.103059] [Medline: 32174347]
21. Durojaiye A, McGeorge N, Puett L, Stewart D, Fackler J, Hoonakker P, et al. Mapping the flow of pediatric trauma patients using process mining. Appl Clin Inform 2018 Jul;9(3):654-666 [FREE Full text] [doi: 10.1055/s-0038-1668089] [Medline: 30134474]
22. Durojaiye A, Levin S, Toerper M, Kharrazi H, Lehmann HP, Gurses AP. Evaluation of multidisciplinary collaboration in pediatric trauma care using EHR data. J Am Med Inform Assoc 2019 Jun 01;26(6):506-515 [FREE Full text] [doi: 10.1093/jamia/ocy184] [Medline: 30889243]
23. Durojaiye A, Puett L, Levin S, Toerper M, McGeorge N, Webster K, et al. Linking electronic health record and trauma registry data: assessing the value of probabilistic linkage. Methods Inf Med 2018 Nov;57(5-06):e3 [FREE Full text] [doi: 10.1055/s-0038-1675220] [Medline: 30453337]
24. Maryland state trauma registry data dictionary for pediatric patients. Maryland Institute for Emergency Medical Services Systems. URL: https://www.miemss.org/home/Portals/0/Docs/OtherPDFs/Web-registry-data-dictionary-pediatric.pdf?ver=2016-03-10-140444-350 [accessed 2022-01-28]
25. Benham-Hutchins MM, Effken JA. Multi-professional patterns and methods of communication during patient handoffs. Int J Med Inform 2010 Apr;79(4):252-267. [doi: 10.1016/j.ijmedinf.2009.12.005] [Medline: 20079686]
26. Aalst W. Process Mining: Discovery, Conformance and Enhancement of Business Processes. Switzerland: Springer; 2011.
27. Rojas E, Munoz-Gama J, Sepúlveda M, Capurro D. Process mining in healthcare: a literature review. J Biomed Inform 2016 Jun;61:224-236 [FREE Full text] [doi: 10.1016/j.jbi.2016.04.007] [Medline: 27109932]
28. van der Aalst WM, Reijers HA, Song M. Discovering social networks from event logs. Comput Supported Coop Work 2005 Oct 28;14(6):549-593. [doi: 10.1007/s10606-005-9005-9]
29. Ferreira D, Cláudia A. Discovering user communities in large event logs. In: Business Process Management Workshops. Berlin, Heidelberg: Springer; 2012.
30. Jenks G. The data model concept in statistical mapping. International Yearbook of Cartography. 1967. URL: https://tinyurl.com/bddrbw69 [accessed 2022-01-28]
31. Brunson J, Laubenbacher RC. Applications of network analysis to routinely collected health care data: a systematic review. J Am Med Inform Assoc 2018 Feb 01;25(2):210-221 [FREE Full text] [doi: 10.1093/jamia/ocx052] [Medline: 29025116]
32. Soulakis N, Carson MB, Lee YJ, Schneider DH, Skeehan CT, Scholtens DM. Visualizing collaborative electronic health record usage for hospitalized patients with heart failure. J Am Med Inform Assoc 2015 Mar;22(2):299-311 [FREE Full text] [doi: 10.1093/jamia/ocu017] [Medline: 25710558]
33. Carson MB, Scholtens DM, Frailey CN, Gravenor SJ, Kricke GE, Soulakis ND. An outcome-weighted network model for characterizing collaboration. PLoS One 2016;11(10):e0163861 [FREE Full text] [doi: 10.1371/journal.pone.0163861] [Medline: 27706199]
34. Landon BE, Keating NL, Barnett ML, Onnela J, Paul S, O'Malley AJ, et al. Variation in patient-sharing networks of physicians across the United States. JAMA 2012 Jul 18;308(3):265-273 [FREE Full text] [doi: 10.1001/jama.2012.7615] [Medline: 22797644]
35. Bholowalia P, Kumar A. EBK-means: A clustering technique based on elbow method and k-means in WSN. Int J Comput Applications 2014;105(9):17-24. [doi: 10.5120/18405-9674]
36. Csardi G, Nepusz T. The igraph software package for complex network research. InterJournal, Complex Systems 2006;1695(5):1-9 [FREE Full text]
37. R: a language and environment for statistical computing. The R Foundation. URL: https://www.R-project.org/ [accessed 2022-01-28]

38. Kalinka A, Tomancak P. linkcomm: an R package for the generation, visualization, and analysis of link communities in networks of arbitrary size and type. Bioinformatics 2011 Jul 15;27(14):2011-2012 [FREE Full text] [doi: 10.1093/bioinformatics/btr311] [Medline: 21596792]
39. Radicchi F, Castellano C, Cecconi F, Loreto V, Parisi D. Defining and identifying communities in networks. Proc Natl Acad Sci U S A 2004 Mar 02;101(9):2658-2663 [FREE Full text] [doi: 10.1073/pnas.0400054101] [Medline: 14981240]
40. Ahn Y, Bagrow JP, Lehmann S. Link communities reveal multiscale complexity in networks. Nature 2010 Aug 05;466(7307):761-764. [doi: 10.1038/nature09182] [Medline: 20562860]
41. Xie J, Kelley S, Szymanski BK. Overlapping community detection in networks. ACM Comput Surv 2013 Aug;45(4):1-35. [doi: 10.1145/2501654.2501657]
42. Liu W, Jiang X, Pellegrini M, Wang X. Discovering communities in complex networks by edge label propagation. Sci Rep 2016 Mar 01;6:22470 [FREE Full text] [doi: 10.1038/srep22470] [Medline: 26926830]
43. McQuitty LL. Hierarchical linkage analysis for the isolation of types. Educ Psychol Measur 2016 Jul 02;20(1):55-67. [doi: 10.1177/001316446002000106]
44. Stata Corp. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP; 2013.
45. McGeorge NM, Durojaiye AB, Fackler JC, Stewart D, Puett LL, Gurses AP. Supporting clinicians teamwork through better HIT design: understanding the trauma care system at a level 1 pediatric trauma center using a mixed methods approach. Pediatric Trauma Society. 2017. URL: https://pediatrictraumasociety.org/meeting/abstracts/2017/P38.cgi [accessed 2018-07-28]
46. Kamada T, Kawai S. An algorithm for drawing general undirected graphs. Inf Processing Letters 1989 Apr;31(1):7-15. [doi: 10.1016/0020-0190(89)90102-6]
47. Adai AT, Date SV, Wieland S, Marcotte EM. LGL: creating a map of protein function with an algorithm for visualizing very large biological networks. J Mol Biol 2004 Jun 25;340(1):179-190. [doi: 10.1016/j.jmb.2004.04.047] [Medline: 15184029]
48. Fruchterman TM, Reingold EM. Graph drawing by force-directed placement. Softw Pract Exper 1991 Nov;21(11):1129-1164. [doi: 10.1002/spe.4380211102]
49. Groh EM, Feingold PL, Hashimoto B, McDuffie LA, Markel TA. Temporal variations in pediatric trauma: rationale for altered resource utilization. Am Surg 2018 Jun 01;84(6):813-819. [Medline: 29981607]
50. Johnson A, Sharma R, Anis J, Kaminski S, Kanard R. Towards a more intelligent model of providing pediatric trauma care: identifying temporal variations in trauma team activations. Trauma Surg Acute Care Open 2020;5(1):e000448 [FREE Full text] [doi: 10.1136/tsaco-2020-000448] [Medline: 33225069]
51. Carmody IC, Romero J, Velmahos GC. Day for night: should we staff a trauma center like a nightclub? Am Surg 2002 Dec;68(12):1048-1051. [Medline: 12516806]
52. Egol KA, Tolisano AM, Spratt KF, Koval KJ. Mortality rates following trauma: the difference is night and day. J Emerg Trauma Shock 2011 Apr;4(2):178-183 [FREE Full text] [doi: 10.4103/0974-2700.82202] [Medline: 21769202]
53. Rucker D. Using telephony data to facilitate discovery of clinical workflows. Appl Clin Inform 2017 Apr 19;8(2):381-395 [FREE Full text] [doi: 10.4338/ACI-2016-11-RA-0191] [Medline: 28421225]

Abbreviations

ADT: admission, discharge, and transfer
ED: emergency department
EHR: electronic health record
HCP: health care professional
IDC: imaging data coordinator
LOWESS: Locally Weighted Scatterplot Smoothing
MTS: multi-team system
PICU: pediatric intensive care unit

Edited by R Kukafka; submitted 16.05.21; peer-reviewed by K Durant, R Xu, A Staffini, A Sadatmoosavi; comments to author 04.07.21; revised version received 30.10.21; accepted 15.11.21; published 04.02.22.

Please cite as:
Durojaiye A, Fackler J, McGeorge N, Webster K, Kharrazi H, Gurses A
Examining Diurnal Differences in Multidisciplinary Care Teams at a Pediatric Trauma Center Using Electronic Health Record Data: Social Network Analysis
J Med Internet Res 2022;24(2):e30351
URL: https://www.jmir.org/2022/2/e30351
doi:10.2196/30351
PMID:35119372

Original Paper

Effectiveness, Cost-effectiveness, and Cost-Utility of a Digital Alcohol Moderation Intervention for Cancer Survivors: Health Economic Evaluation and Outcomes of a Pragmatic Randomized Controlled Trial

Ajla Mujcic[1,2], MSc; Matthijs Blankers[2,3,4], PhD; Brigitte Boon[5,6,7], PhD; Anne H Berman[8,9,10], PhD; Heleen Riper[11,12,13], PhD; Margriet van Laar[2], PhD; Rutger Engels[1], PhD

[1]Erasmus School of Social and Behavioural Sciences, Erasmus University Rotterdam, Rotterdam, Netherlands

[2]Trimbos Institute, Utrecht, Netherlands

[3]Department of Psychiatry, Amsterdam University Medical Center, Location Amsterdam Medical Center, University of Amsterdam, Amsterdam, Netherlands

[4]Department of Research, Arkin Mental Health Care, Amsterdam, Netherlands

[5]Academy het Dorp, Arnhem, Netherlands

[6]Siza, Arnhem, Netherlands

[7]Tranzo, Tilburg University, Tilburg, Netherlands

[8]Department of Psychology, Uppsala University, Uppsala, Sweden

[9]Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institute, Stockholm, Sweden

[10]Stockholm Health Care Services, Stockholm Region, Stockholm, Sweden

[11]Section Clinical Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

[12]Department of Psychiatry, Medical University, University of Turku, Turku, Finland

[13]Department of Psychiatry, Amsterdam Public Health Research Institute, Amsterdam University Medical Center, Amsterdam, Netherlands

Corresponding Author:
Ajla Mujcic, MSc
Erasmus School of Social and Behavioural Sciences
Erasmus University Rotterdam
Burgemeester Oudlaan 50
Rotterdam, 3062 PA
Netherlands
Phone: 31 30 29 59 256
Email: amujcic@trimbos.nl

Abstract

Background: Alcohol moderation (AM) interventions may contribute to better treatment outcomes and the general well-being of cancer survivors.

Objective: This study evaluates the effectiveness, cost-effectiveness, and cost-utility of MyCourse, a digital AM intervention, compared with a noninteractive digital information brochure for cancer survivors.

Methods: A health economic evaluation alongside a pragmatic 2-arm parallel-group randomized controlled trial was conducted with follow-ups at 3, 6, and 12 months after randomization. The study was conducted on the web in the Netherlands from 2016 to 2019. Participants were adult 10-year cancer survivors drinking over the Dutch-recommended drinking guidelines (≤7 standard units [10 g of alcohol] per week) with the intention to moderate or quit drinking. Overall, 103 participants were randomized and analyzed: 53 (51.5%) in the MyCourse group and 50 (48.5%) in the control group. In the MyCourse group, participants had access to a newly developed, digital, minimally guided AM intervention, *MyCourse–Moderate Drinking*. The primary outcome was the self-reported number of standard drinks (10 g of ethanol) consumed in the past 7 days at the 6-month follow-up. The secondary outcome measures were alcohol-related problems as measured by the Alcohol Use Disorders Identification Test (AUDIT) and treatment satisfaction. For the health economic evaluation, health care costs, costs because of productivity losses, and intervention costs were assessed over a 12-month horizon.

Results: Alcohol use at the 6-month follow-up decreased by 38% in the MyCourse group and by 33% in the control group. No difference in 7-day alcohol use was found between the groups (*B*=2.1, 95% CI −7.6 to 3.1; *P*=.22) at any of the follow-ups. AUDIT scores for alcohol-related problems decreased over time in both groups, showing no significant difference between the groups (Cohen *d*=0.3, 95% CI −0.1 to 0.6; *P*=.21). Intervention costs per participant were estimated at US $279 for the MyCourse group and US $74 for the control group. The mean societal costs were US $18,092 (SD 25,662) and US $23,496 (SD 34,327), respectively. The MyCourse group led to fewer gained quality-adjusted life years at lower societal costs in the cost-utility analysis. In the cost-effectiveness analysis, the MyCourse group led to a larger reduction in drinking units over time at lower societal costs (incremental cost-effectiveness ratio per reduced drink: US $ −1158, 95% CI −1609 to −781).

Conclusions: At 6 months, alcohol use was reduced by approximately one-third in both groups, with no significant differences between the digital intervention MyCourse and a noninteractive web-based brochure. At 12 months, cost-effectiveness analyses showed that MyCourse led to a larger reduction in drinking units over time, at lower societal costs. The MyCourse group led to marginally fewer gained quality-adjusted life years, also at lower societal costs.

Trial Registration: Netherlands Trial Register NTR6010; https://www.trialregister.nl/trial/5433

International Registered Report Identifier (IRRID): RR2-10.1186/s12885-018-4206-z

(J Med Internet Res 2022;24(2):e30095) doi:10.2196/30095

KEYWORDS

alcohol; brief interventions; cancer survivors; effectiveness; cost-effectiveness; eHealth; mobile phone

Introduction

Alcohol use is one of the main lifestyle factors influencing cancer development, and there is also evidence that it negatively affects the development of new malignancies [1], cancer treatment success [2], and mortality rates [3]. Therefore, it is recommended that cancer survivors quit or minimize alcohol use [4]. Currently, drinking rates among cancer survivors are comparable with drinking rates among the general population, with estimates that 5.1% of cancer survivors are heavy drinkers (>2 drinks per day for men; >1 drink per day for women) versus 6% of the general population [5].

Studies evaluating alcohol moderation (AM) interventions in cancer survivors are scarce. AM interventions offer support in reducing or quitting alcohol use and can range from brief face-to-face interventions by health care providers to smartphone app–based interventions. A 2018 review on smoking and alcohol cessation interventions in patients with head and neck cancer and oral dysplasia [6] found no randomized controlled trials (RCTs) evaluating AM interventions for head and neck cancer survivors, and neither did a recent meta-analysis on AM distance–based interventions for cancer survivors of all cancer types [7]. The latter review identified a few studies that incorporated AM as a module in a broader lifestyle program and found insufficient evidence of the interventions' effectiveness on AM. A qualitative study assessed patients' experiences with a face-to-face alcohol cessation program in bladder cancer survivors undergoing surgery; results indicated that major bladder surgery was a useful cue for motivating patients with cancer to reconsider the consequences of risky drinking, and the alcohol intervention was seen as a relevant offer around the time of surgery [8]. Facilitating access to AM interventions for cancer survivors via distance-based interventions, particularly digital ones, might be an effective and highly accessible means to provide the growing population of cancer survivors with AM support [9].

Studies among the general population have shown that brief face-to-face and digital interventions can be effective in reducing alcohol consumption. An individual patient data meta-analysis comparing guided and unguided low‐intensity internet interventions for AM found that participants in both types of interventions used on average 50 g less ethanol per week than the controls (5.02 standard units of 10 g of ethanol, 95% CI −7.57 to −2.48) [10]. A conventional meta-analysis evaluating brief AM interventions delivered in a primary care setting found that participants used on average 20 g (95% CI −28 to −12) less ethanol per week than the controls [11]. A meta-analysis of personalized digital interventions found similar results when comparing the interventions to nonintervention control groups (23 g less ethanol per week in participants receiving a digital intervention compared with no or minimal interventions, 95% CI 15-30 based on 41 studies) and found no difference in reduction of alcohol consumption in personalized digital interventions compared with face-to-face interventions, based on 5 studies [12].

Brief alcohol interventions in primary care settings have been found to be cost-effective for the general population [13]. Referral to a digital AM intervention was a cost-effective strategy in 3 European countries [14]. A game with tailored feedback on alcohol awareness was found to be cost-effective from the societal perspective in reducing the number of drinks in subgroups (older age and lower educational level) of adolescents [15]. We found no studies on the cost-effectiveness of digital AM interventions for cancer survivors, but cost-effectiveness is a key element in the knowledge base needed for policy decisions regarding implementation and financing of digital interventions [16,17]. It is unknown what results in effectiveness and cost-effectiveness are to be obtained from a digital AM intervention that is tailored to cancer survivors, as cancer survivors have increased feelings of distress and symptoms of anxiety and depression [18,19], and they could have additional benefits of AM (eg, in terms of treatment outcomes) [2].

Therefore, it was deemed necessary to evaluate both the effectiveness and cost-effectiveness of a minimally-guided digital intervention aimed at supporting cancer survivors to moderate their alcohol use: *MyCourse–Moderate Drinking* (in Dutch: *MijnKoers–Minderen met Drinken*). The development process and a detailed intervention description are provided elsewhere [20]. In this study, we aim to answer the following research questions: (1) Is the digital, minimally guided AM intervention *MyCourse–Moderate Drinking* more effective than a digital AM brochure to moderate alcohol use? (2) From a societal perspective, is the digital, minimally-guided AM intervention *MyCourse–Moderate Drinking* more cost-effective than a web-based AM brochure in terms of incremental costs per reduced weekly drink and incremental costs per quality-adjusted life year (QALY) gained?

We expect the MyCourse intervention to be both more effective and more cost-effective than a web-based brochure on AM.

Methods

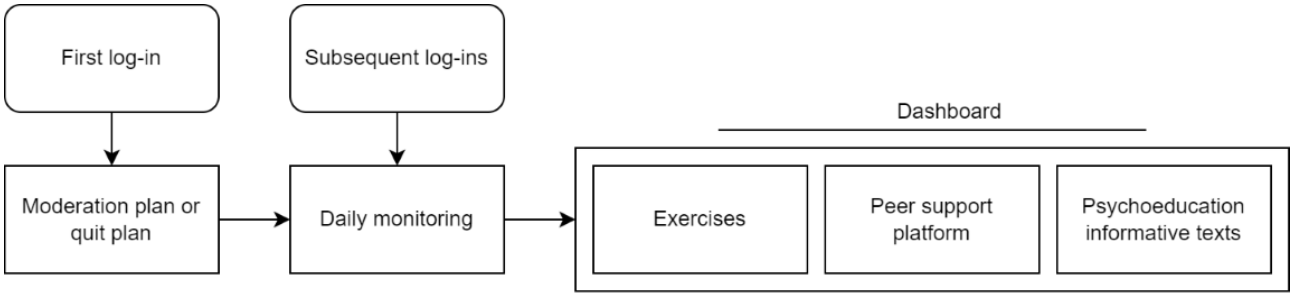
Design

In a 2-arm, individually randomized RCT conducted in the Netherlands between 2016 and 2019, the effectiveness and cost-effectiveness of the *MyCourse–Moderate Drinking* intervention for cancer survivors was evaluated. The first inclusion was on November 28, 2016, and the last inclusion was on September 3, 2018. The last follow-up measurement was collected on September 30, 2019. The study was prospectively registered in the Netherlands Trial Register (NTR6010). The planned inclusion period was extended by several months to recruit as many participants as possible. An extensive description of the study protocol has been provided in the study by Mujcic et al [20]. This study was part of a set of 2 separate RCTs on digital interventions for AM and smoking cessation in cancer survivors. The results of the RCT on the smoking cessation intervention (*MyCourse–Quit Smoking*) will be published separately. Ethical approval was obtained from an accredited medical research and ethics committee in the Netherlands (Toetsingscommissie Wetenschappelijk Onderzoek Rotterdam NL55921.101.16).

Participants and Recruitment

Participants could find out about the study and apply for participation via a web-based screening questionnaire on a dedicated website that was created for the study. Participants were eligible if they were aged ≥18 years, had been diagnosed with any form of cancer in the past 10 years, had a PC or laptop and internet connection at home, had the ability and intention to participate in the study for 12 months, used alcohol more than recommended by Dutch guidelines (operationalized as drinking >7 European standard units of alcohol [70 g of ethanol] per week), and had the intention to reduce their alcohol use. The exclusion criteria were insufficient mastery of the Dutch language, pregnancy or self-reported suicidal ideation, acute psychosis, severe alcohol dependence, dementia, or severe depression at the time of screening. The same screening questionnaire procedure was used for both this trial and the similar parallel trial evaluating a smoking cessation intervention [20]. Some people were eligible for both trials; if so, they were allowed to participate in only one of the trials, based on their own choice (Figure 1 [20]).

Figure 1. Intervention flowchart (adapted from Mujcic et al [20]).



Both web-based and offline strategies were used for the recruitment of participants. Targeted web-based advertisements on social and other media, and on search engines referred those interested to the website and the web-based screening questionnaire. Patient organizations, oncology departments in hospitals, and meeting centers for cancer survivors were contacted and offered promotional material (flyers and posters) to help refer cancer survivors to the website.

Procedure

After filling out the screening questionnaire on the study website, applicants were informed by a computer-generated email about their eligibility for study participation. Those eligible were sent an invitation email containing all relevant patient information, the informed consent form, and a link to register. Eligible cancer survivors had up to 30 days to decide about their participation, and during this period, they could contact the research team or an independent physician with questions. After they had digitally signed the informed consent form, they were sent the baseline questionnaire. Immediately after completion of the baseline measurement, participants were automatically allocated to either the MyCourse or the control group arm in a 1:1 ratio through adaptive randomization (minimization of baseline imbalances with regard to age, sex, and education level) through a server-sided hypertext preprocessor script using a Mersenne twister random number generator. Participants received an email confirming their allocation and containing their username and instructions on how to log on. They were not blinded to study condition allocation (the participants were not explicitly informed about their allocation, but recruitment material included a video showing interactive elements of MyCourse, making it plausible that the participants knew they were not allocated to the

experimental condition. Thus, we cannot consider the participants blinded). At 3, 6, and 12 months after randomization, the participants received a link to the web-based questionnaire via email. The nonresponders received up to 3 reminder emails, and in case of continued nonresponse they were contacted by telephone. For each completed follow-up assessment, they were reimbursed with €25 (approximately US $30). As this was a pragmatic RCT, patients in both groups were not asked to refrain from using additional support if they wanted to.

MyCourse Group

MyCourse–Moderate Drinking is a newly developed, minimally-guided, digital intervention aimed at supporting AM in cancer survivors. It is based on well-established therapeutic approaches: motivational interviewing, cognitive behavioral therapy, and acceptance and commitment therapy, as well as a Dutch digital AM intervention previously found to be effective in the general population [21]. Throughout the development process, cancer survivors and professional experts in eHealth, oncology, and substance use disorders were involved through a series of interviews and focus groups. The intervention was accessible through PC, tablet, and smartphone. At first log-in, the participants were guided by website prompts in either setting up a quit plan or a moderation plan, including a quit date or moderation date, after which they would gain access to several exercises, a web-based diary for self-monitoring of alcohol use and contextual cues, and a peer support platform (See Figure 1 adapted from Mujcic et al [20] and Multimedia Appendix 1 [22]). MyCourse could be used by the participants whenever they chose to, but they were encouraged to log in daily for at least 4 weeks. The intervention and its development process have been extensively described elsewhere [20].

Control Group

The control group consisted of a noninteractive web-based static information brochure on the risks of (increased) alcohol use and tips on how to moderate or quit drinking. It was accessible to the participants at any time by logging into the website. The brochure contained both general information on AM and information specifically relevant to cancer survivors. However, no interactive elements of the MyCourse condition were present, and the participants did not receive reminders.

Additional Support

Both groups were provided with the contact details of the national AM information line (in Dutch: *Alcoholinfo-lijn*), which could help refer participants to additional support if they deemed the received intervention to be insufficient. At the end of the study, at 12 months after randomization, the participants in the control group received access to the digital intervention, *MyCourse–Moderate Drinking*, which was offered to the MyCourse group.

Measures

Baseline

At baseline, we assessed the sociodemographic characteristics and type of cancer. Alcohol use was assessed using Timeline Followback (TLFB) self-reports [23] (number of standard drinks consumed in the past 7 days, ie, 7-day alcohol use). Problematic alcohol use was assessed using the 10-item Alcohol Use Disorders Identification Test (AUDIT) questionnaire. In participants reporting smoking, we assessed tobacco use with TLFB self-reports and nicotine dependence using the 6-item Fagerstrom Nicotine Dependence Test questionnaire [24]. Socially desirable answering tendencies, which may have affected the reliability of the self-reported questionnaire data, were assessed using the Marlowe-Crowne Social Desirability Scale (MCSDS) [25]. We used the 5-level EuroQol 5 Dimension (EQ-5D-5L) [26] measure to assess QALYs. The Medical Outcomes Study Short Form [27] was used to calculate the Short Form 6-dimension (SF-6D) quality of life measure using the Brazier algorithm [28].

Follow-up Measurements

At all follow-up measurements, we assessed alcohol use with TLFB self-reports, quality of life using the EQ-5D-5L and the Medical Outcomes Study Short Form, productivity and health care costs, and use of other AM support. Intervention use variables (eg, number of log-ins and use of major content elements) were collected automatically. The AUDIT questionnaire was administered at the 6- and 12-month follow-ups. At the 3-month follow-up, treatment satisfaction was assessed using a Dutch translation of the German adapted Client Satisfaction Questionnaire (Fragebogen zur Messung der Patientenzufriedenheit; Patient Satisfaction Questionnaire [ZUF-8]) [29]. The use of additional support for AM was retrospectively assessed at follow-up.

Primary and Secondary Outcome Measures

The primary outcome measure was 7-day alcohol use (number of standard drinks; 1 standard drink=10 g of ethanol) at the 6-month follow-up measured by TLFB self-reports. The 6-month assessment was the primary end point, as the studies that formed the basis of our power analysis were based on outcomes at the 6-month follow-up and it is a common end point in alcohol trials [30]. Those who reported no drinking at all in the past 7 days were considered abstinent (score: 0/1). Secondary outcome measures were AUDIT problematic alcohol use (score: 0-40), ZUF-8 treatment satisfaction (score: 8-32), EQ-5D-5L quality of life (score: 0-1), health care costs, and productivity loss.

Costs

Costs were calculated from a societal perspective for the index year 2019. Intervention costs included intervention depreciation costs, costs for hosting the website, technical support, and recruitment costs (which consisted of both advertising costs in web-based and offline media as well as printing costs of promotional material). Recruitment costs were included as they were considered an essential part of the MyCourse and control groups. Health care costs were calculated by multiplying the number of reported contacts with a health care professional with the standard unit cost prices for the Netherlands [31]. Health service costs stemmed from contacts with specialized somatic and mental health care, plus the patients' out-of-pocket costs for home care, but travel costs were not included because, in both groups, the interventions were delivered over the internet. Other health care costs included appointments for physiotherapy,

alternative medicine, and social work. Medication costs were calculated by multiplying the reported dose of a drug with its unit cost price [32].

Productivity loss included costs from absenteeism and presenteeism, calculated according to the friction cost method, meaning productivity losses were limited to a maximum of 85 days, after which production losses cease to exist because the sick employee has been replaced by another and calculated using an elasticity factor of 0.8 as there is not a strict 1:1 relation between days not worked and productivity losses. Cost data related to health care use and productivity loss were assessed using the Trimbos/Institute for Medical Technology Assessment questionnaire for costs associated with psychiatric illness [33] at all follow-up assessments. Cumulative societal costs over the entire follow-up period of 12 months were calculated from the sum of health care costs and productivity losses. Costs were converted from euros to US dollars using purchasing power parities for the reference year 2019. No discount rate was applied as the follow-up period was 12 months.

Sample Size

The sample size was based on conventional levels of statistical significance ($\alpha \leq .05$). On the basis of the average of 2 previous RCTs on very similar self-help interventions in the Netherlands versus a control group [21,30,34], a Cohen *d* effect size of 0.40 was expected. Using the power calculation package pwr for R 3.0.1 (R Foundation for Statistical Computing) [35], we calculated that a sample size of 2 × 57 participants in the case of 1-sided testing led to a power of 0.77 or a power of 0.66 in the case of 2-sided testing. The choice for 1-sided testing was discussed in the previously published protocol paper [20].

Statistical Analyses

Imputation of Missing Data

Except for the ZUF-8 (treatment satisfaction) questionnaire, all primary and secondary outcome measures were analyzed in accordance with the intention-to-treat principle. To that end, missing data for primary and secondary outcome measures, and costs, were multiple-imputed using the predictive mean matching method from the mice package in R [36]. For each missing observation, 50 imputations were created. The responses to the ZUF-8 questionnaire were not imputed.

Effect Evaluation

Alcohol use in the past 7 days (count data 0, 1...,N) was analyzed using robust estimation of generalized linear mixed models from the robustlmm package in R [37], as the data did not fit well into any of the commonly supported distributions. Imputation of missing values before running a generalized linear mixed model allowed us to consider all variables that could have affected the dropout. Covariates in the model were the minimized variables (gender, age, and education) and the MCSDS (to statistically account for any social desirability of responses). Model estimates, Cohen *d*, 95% CIs, and *P* values were reported. Differences over time and between the groups on AUDIT problematic alcohol use and ZUF-8 patient satisfaction scores were analyzed using a linear mixed model in the lme4 package in R [38]. We used 1-sided testing and an α of .05 as described in the study protocol [20].

Cost-effectiveness Analyses

An economic evaluation was conducted alongside the RCT in concordance with the Consolidated Health Economic Evaluation Reporting Standards Statement [39] and following the approach by Drummond et al [40]. QALYs over the entire follow-up period were computed using the Dutch tariff (utility weights) [41] and the area under the curve method. The incremental cost-effectiveness ratio (ICER) was calculated as follows:

$$ICER = (C_1 - C_0) / (E_1 - E_0) \quad \textbf{(1)}$$

where C refers to costs, E refers to effects, and the subscripts 0 and 1 refer to the MyCourse and control arms, respectively. We generated 2500 nonparametric bootstrapped samples and plotted the corresponding incremental costs and incremental effects on a cost-effectiveness plane. Both ICER per QALY and ICER per reduced weekly drink were calculated from the following 4 perspectives: societal, health care, productivity loss, and intervention cost only. Cost-effectiveness acceptability curves (CEACs) were also drawn to assess the likelihood that the experimental intervention will be deemed cost-effective given a series of willingness-to-pay ceilings.

Sensitivity Analyses

The robust regression on the mice-imputed data was the main analysis. We conducted several sensitivity analyses to assess effectiveness and cost-effectiveness using QALYs based on the SF-6D (instead of the EQ-5D-5L), imputation using the Amelia II package instead of the mice package, Winsorization of costs, and different statistical models (Multimedia Appendix 1).

Results

Sample Characteristics

The participant flow and retention rates are shown in Figure 2. Of the 2346 ineligible people, 1684 (71.78%) had had no diagnosis of cancer in the past 10 years. A total of 321 cancer survivors were eligible for participation in the study, of whom 206 (64.2%) declined to participate, and 34 (10.6%) participated in the study on smoking cessation instead. Of 115, 10 (8.7%) cancer survivors did not complete the baseline questionnaire and were therefore not randomized, and 2 (1.7%) cancer survivors withdrew during the course of the study. This resulted in a study sample of 103 participants; of whom, 53 (51.5%) were randomized into the experimental MyCourse group and 50 (48.5%) were randomized into the control group. Table 1 presents the sociodemographic and other characteristics of the sample. In summary, the mean age of the participants was 54.6 (SD 11) years, 16.5% (17/103) were men, most were married or living with a significant other (70/103, 68%), and 31.3% (32/103) had a middle or low educational level. Breast cancer was the most frequently reported type of cancer (65/103, 63.1%). Problematic alcohol use as measured by the AUDIT was significantly higher in the MyCourse group than in the control group as calculated using the Welch 2-sample *t* test (2-tailed) for continuous variables ($t_{100.9}$=2.03; *P*=.02). No difference was found in the proportion of missing data between the groups

(χ^2_1=2.5 *P*=.11; see Multimedia Appendix 2 for details). Data were missing because of loss to follow-up (participants who did not respond after several reminders by email and telephone).

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) flowchart. RCT: randomized controlled trial.

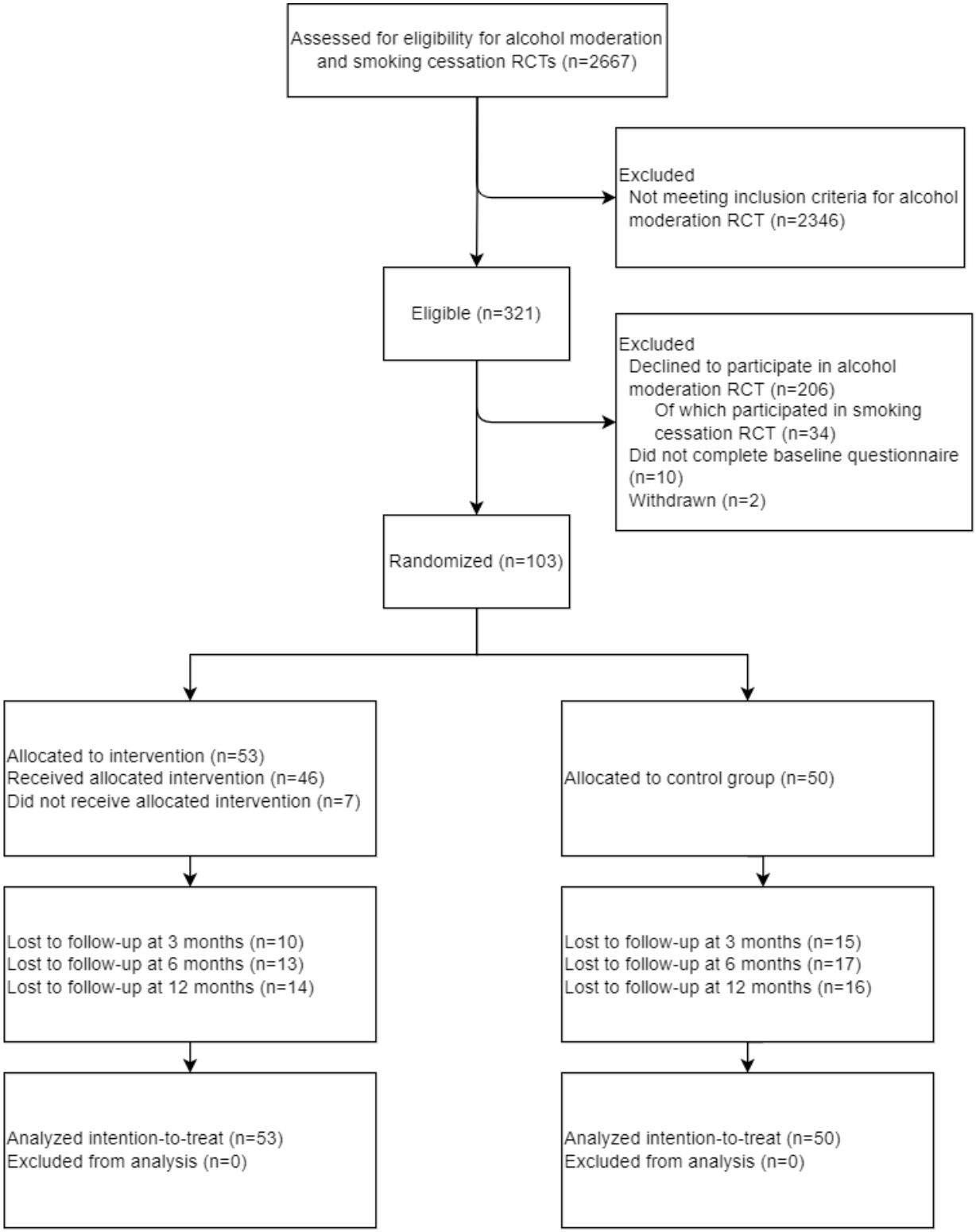


Table 1. Baseline characteristics[a].

Characteristic	MyCourse (n=53)	Control (n=50)	Total (N=103)
Gender, n (%)			
Women	46 (87)	40 (80)	86 (83)
Men	7 (13)	10 (20)	17 (17)
Age (years), mean (SD)	54.5 (12.1)	54.6 (9.9)	54.6 (11)
Education, n (%)			
Higher level	34 (64)	37 (74)	71 (69)
Midlevel	11 (21)	11 (22)	22 (21)
Lower level	8 (15)	2 (4)	10 (10)
Marital status, n (%)			
Married or living together	34 (64)	36 (72)	70 (68)
Unmarried or living alone	6 (11)	8 (16)	14 (14)
Divorced	10 (19)	4 (8)	14 (14)
Widowed	3 (6)	2 (4)	5 (5)
Drinking behavior, mean (SD)			
Number of drinks in past 7 days	26.8 (19.0)	20.7 (14.7)	23.8 (17.2)
AUDIT[b]	14.5 (6.0)	12.2 (5.4)	13.3 (5.8)
Smoking behavior			
Smoked in last month, n (%)	10 (19)	6 (12)	16 (16)
Number of cigarettes in past 7 days among smokers, mean (SD)	87.9 (52.6)	81.6 (68.5)	85.3 (56.8)
Nicotine dependence, mean (SD)	0.6 (1.7)	0.3 (1.3)	0.5 (1.5)
Cancer diagnosis, n (%)			
Breast	38 (72)	27 (54)	65 (63)
Uterus	4 (8)	2 (4)	6 (6)
Head and neck	1 (2)	4 (8)	5 (5)
Colon	2 (4)	3 (6)	5 (5)
Lung	1 (2)	2 (4)	3 (3)
Other (including bladder, lymphatic, melanoma, skin, and prostate)	7 (13)	12 (24)	19 (18)

[a]Percentages may not add up to 100 because of rounding.

[b]AUDIT: Alcohol Use Disorders Identification Test.

Treatment Uptake and Satisfaction

Overall, patients were most satisfied in the MyCourse group (Cohen d=0.81; $t_{61.5}$=3.42; P<.001; see Multimedia Appendix 2 for the mean scores). Most participants logged in at least once (46/53, 87%). The average number of times the participants logged in was 31.4 (SD 50.5), with a median of 8 (range 0-254). For those who logged in at least once, the period between the first and last log-in was on average 105.6 (SD 125.6) days with a median of 45 days. There was little use of AM support besides MyCourse; no support was reported most often (control group: 26/50, 52%; MyCourse group: 26/53, 49%) and some connected with others who were also moderating their drinking (control group: 4/50, 8%; MyCourse group: 5/53, 9%). Of 103 participants, only 1 (0.9%) reported having had contact with a health care professional about AM.

Incremental Effects

Primary Outcome

Despite the randomization, there was an apparent, although nonsignificant, difference between the groups in baseline alcohol use (Table 1). The number of drinks consumed in the past week at the 6-month follow-up decreased by 38% in the MyCourse group and by 33% in the control group and even more at the 12-month follow-up—by 48% in the MyCourse group and 38% in the control group (Table 2). No difference in 7-day alcohol use was found between the groups (unstandardized regression coefficient, B=−2.1, 95% CI −7.6 to 3.1; P=.22; Table 3) at 6 months—or at any of the other follow-up assessments—when controlling for MCSDS score, baseline alcohol use, gender, age, and education.

Table 2. Drinking behavior outcomes at the 3-, 6-, and 12-month follow-ups (N=103)[a].

Variable	MyCourse (n=53)	Control (n=50)
Number of drinks in past 7 days, mean (SD)[b]		
Baseline	26.8 (19.0)	20.7 (14.7)
3-month follow-up	17.3 (15.8)	15.1 (11.9)
6-month follow-up	16.6 (15.2)	13.8 (11.4)
12-month follow-up	13.9 (11.0)	12.9 (10.7)
Change in number of drinks in past 7 days, mean (SD)[c]		
3-month follow-up	−8.5 (12.0)	−5.2 (13.5)
6-month follow-up	−9.4 (15.0)	−6.4 (16.4)
12-month follow-up	−12.1 (16.3)	−7.4 (13.3)
AUDIT[d], mean (SD)		
Baseline	14.5 (6.0)	12.2 (5.4)
6-month follow-up	11.3 (6.2)	9.9 (5.1)
12-month follow-up	10.0 (6.0)	9.3 (5.1)
Abstinence, n (%)		
3-month follow-up	6 (11)	7 (14)
6-month follow-up	6 (11)	8 (16)
12-month follow-up	5 (10)	7 (14)

[a]Missing data were imputed.

[b]The number of drinks per day was maximized at 11 units in the follow-up measurements for the imputation of missing data, meaning that 77 was the maximum number of drinks in the past 7 days.

[c]Mean number of drinks at follow-up minus the mean number of drinks at baseline.

[d]AUDIT: Alcohol Use Disorders Identification Test.

Table 3. Treatment effects on drinking behavior at the 3-, 6-, and 12-month follow-ups[a].

Outcome measure	Treatment effect		
	$B_{adjusted}$ (SE; 95% CI)	*P* value	Cohen *d* (95% CI)
Number of drinks in past 7 days[b]			
3-month follow-up	−3.2 (2.6; −8.5 to 1.9)	.11	N/A[c]
6-month follow-up	−2.1 (2.7; −7.6 to 3.1)	.22	N/A
12-month follow-up	−3.7 (2.7; −8.9 to 1.6)	.09	N/A
AUDIT[d]			
6-month follow-up	−0.9 (1.0)[e]	.21	0.3 (−0.1 to 0.6)
12-month follow-up	−1.6 (1.0)[e]	.06	0.1 (−0.2 to 0.5)

[a]Missing data were imputed.

[b]Adjusted coefficients are based on a robust regression mixed model with random intercept and fixed slope in which the outcome measure at follow-up is regressed upon baseline number of drinks, covariates, and group.

[c]N/A: not applicable.

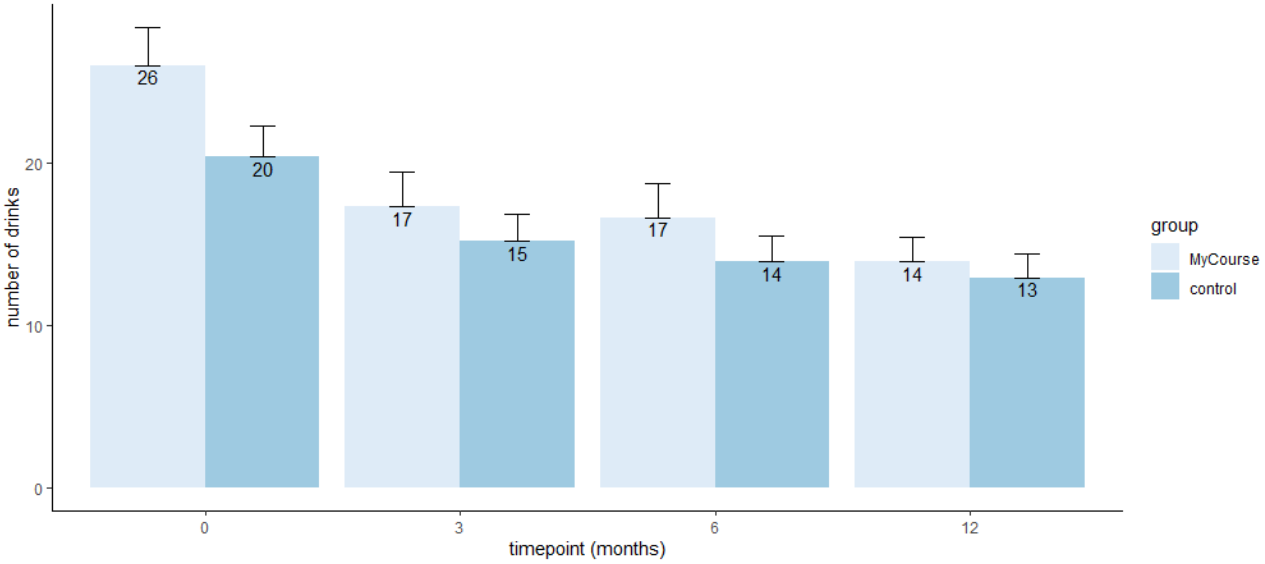
[d]AUDIT: Alcohol Use Disorders Identification Test (adjusted coefficients are based on a linear mixed model with random intercept and fixed slope in which the outcome measure at follow-up is regressed upon baseline number of drinks, covariates, and group).

[e]95% CI value is not available.

Secondary Outcomes

AUDIT scores decreased over time in both groups (Table 2 and Figure 3), but there was no difference between the groups (Cohen *d*=0.3, 95% CI −0.1 to 0.6; *P*=.21; Table 3) at 6 months. The mean EQ-5D-5L QALYs score in the MyCourse group was 0.82 (SD 0.12) and in the control group 0.84 (SD 0.10). There was no significant effect of the treatment on the quality of life based on EQ-5D-5L scores ($B_{adjusted}$=0.003, SE 0.01; *P*=.39).

Figure 3. Mean number of drinks in the past 7 days in both groups at baseline and during the course of the study, including SEs.



Incremental Costs

Table 4 presents the costs per group and the incremental costs (cost difference between the MyCourse and control groups) per cost item. The intervention was costed at US $279 per participant in the MyCourse group and US $74 per participant in the control group. The average health care costs accumulated over the full 12-month follow-up time were US $7840 (SD 11,767) per participant in the MyCourse group and US $8233 (SD 15,077) per participant in the control group, and the incremental health care costs were US $−393. Costs owing to productivity losses were mainly driven by absenteeism: US $9532 (SD 19,389) per participant in the MyCourse group and US $14,799 (SD 23,364) per participant in the control group, with high within-group variance. Incremental productivity costs per participant were on average US $−5217 (SD 26,378). The average cumulative societal costs were US $5404 (SD 42,859) lower in the MyCourse group compared with the control group.

Table 4. Mean cumulative costs (in US $) by group and incremental costs (N=103).

Cost item	MyCourse (n=53), mean (SD)	Control (n=50), mean (SD)	Incremental costs[a] (n=53), mean (SD)
Health care costs	7840 (11,767)	8233 (15,077)	−393 (19,125)
Specialized somatic	3819 (5772)	3627 (6463)	192 (8665)
Specialized psychiatric	1209 (3878)	688 (1906)	521 (4321)
Patient and family costs	953 (5517)	178 (1811)	775 (5807)
Other	907 (1142)	1126 (1434)	−219 (1833)
Medication	953 (5479)	2613 (10,521)	−1660 (11,862)
Productivity loss	9972 (1934)	15,189 (26,307)	−5217 (26,378)
Presenteeism	153 (319)	210 (408)	−57 (518)
Absenteeism	9532 (19,389)	14,799 (26,364)	−5267 (32,726)
Unpaid work	452 (1049)	474 (1007)	−22 (1454)
Intervention costs	279 (0)	74 (0)	205 (0)
Total societal costs	18,092 (25,662)	23,496 (34,327)	−5404 (42,859)

[a]Costs in the MyCourse group minus costs in the control group.

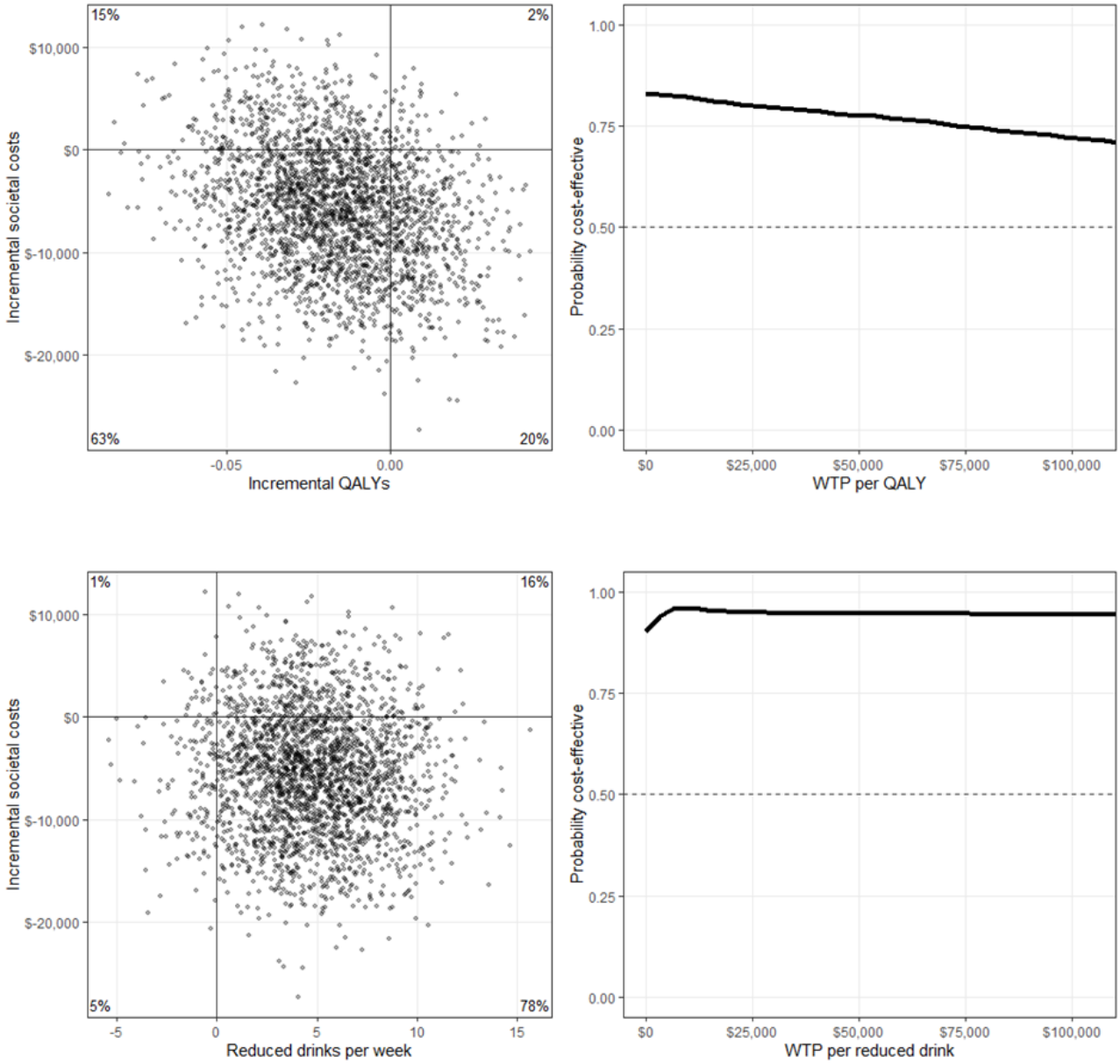
Cost-Utility

With QALY as the outcome, the ICER was US $314,606 (95% CI 186,201-553,552). The cost-effectiveness plane (Figure 4) shows that there is a 63% chance that the MyCourse intervention will lead to fewer QALYs gained at lower societal costs and a 15% chance that it will lead to more QALYs gained at lower societal costs. The relatively high ICERs are mostly a result of

a difference in productivity costs (Table 4) and a very small differential effect on QALYs.

Assuming an intervention-cost-only perspective, the ICER per QALY gained became negative (US $−11,930; 95% CI −18,440 to −8912), indicating that intervention costs were higher and the QALY gains were lower in the MyCourse group compared with the control group.

Figure 4. Cost-effectiveness planes and cost-effectiveness acceptability curves in US $. QALY: quality-adjusted life year; WTP: willingness to pay.



Cost-effectiveness

The MyCourse group reduced their number of weekly drinks more on average (mean 12.1, SD 16.3 drinks) than the control group (mean 7.4, SD 13.3 drinks) over a 12-month period and at lower societal costs. ICER per reduced drink was calculated at US $−1158 (95% CI −1609 to −781), indicating that compared with the control group each additional reduced drink in the MyCourse group was associated with a societal cost reduction. There is a 78% chance that the MyCourse intervention will lead to more weekly reduced drinks at lower societal costs and a 16% chance that it will lead to more weekly reduced drinks at higher costs (Figure 4). MyCourse will be preferred over the control group at any willingness-to-pay level (see Figure 4 for the CEAC curve). Assuming an intervention-cost-only perspective, a reduction of 1 additional weekly drink would cost an additional US $44 (95% CI 38-53) in the MyCourse group compared with the control group. Table 5 shows a breakdown by perspective.

Table 5. Incremental cost-effectiveness ratios between baseline and the 12-month follow-up[a].

Perspective	Incremental costs per QALY[b] (US $)	Incremental costs per reduced drink (US $)
	Value, mean (95% CI)	Value, mean (95% CI)
Health care	22,859 (−18,584 to 78,705)	−84 (−242 to 74)
Productivity loss	303,677 (198,917 to 516,624)	−1118 (−1497 to −823)
Intervention cost only	−11,930 (−18,440 to −8912)	44 (38 to 53)
Societal	314,606 (186,201 to 553,552)	−1158 (−1609 to −781)

[a]The incremental cost-effectiveness ratios were calculated as follows: $(C_1–C_0) / (E_1–E_0)$, where C refers to costs, E refers to effects, and the subscripts 0 and 1 refer to the experimental and control arms, respectively.

[b]QALY: quality-adjusted life year (as measured by the 5-level EuroQol 5 Dimension).

Sensitivity Analyses

Sensitivity analyses of the Amelia II-imputed data ($B_{adjusted}$=2.1, SE 1.8; *P*=.12) and completers-only data ($B_{adjusted}$=1.7, SE 1.4; *P*=.12) as well as a negative binomial mixed model on the mice-imputed data (incidence rate ratio=1.05, 95% CI 0.79-1.4; *P*=.38) corroborated the main findings and showed no effect of treatment on the number of drinks in the past week. All sensitivity analyses showed a decrease in alcohol use in both groups. Because of the apparent, although nonsignificant difference, between the groups in baseline alcohol use, we also modeled the individual change scores in alcohol use in a robust regression. Although the average reduction in alcohol use at 6 and 12 months was larger in the MyCourse group, the difference was not statistically significant, yielding the same results ($B_{adjusted}$=−2.1, SE 1.8; *P*=.12; Table 2). When QALYs were based on SF-6D scores, results of the economic evaluation remained similar. When Winsorization of extreme costs was applied at the 95th percentile, the cost-effectiveness planes and CEAC curves remained similar (Multimedia Appendix 3); however, ICER per EQ-5D-5L QALY was US $118,287 (95% CI 51,324-235,817), and ICER per reduced drink became less extreme (US $−435, 95% CI −680 to −219). Overall, the sensitivity analyses attested to the robustness of the findings in the main analysis.

Discussion

Principal Findings

We evaluated the effectiveness and cost-effectiveness of *MyCourse–Moderate Drinking*, a digital AM intervention for cancer survivors versus a web-based noninteractive information brochure. At the 6-month follow-up, the number of drinks in the past 7 days was reduced by 38% in the MyCourse group (mean −9.4, SD 15.0 standard units) and by 33% in the control group (mean −6.4, SD 16.4 standard units) and even further at the 12-month follow-up (MyCourse group: mean −12.1, SD 16.3 standard units; control group: mean −7.4, SD 13.3 standard units). No significant difference in 7-day alcohol use was found between the groups at any of the follow-up points. AUDIT scores decreased over time in both groups, but there was no statistically significant difference between the MyCourse group and the control group. Importantly, the participants were more satisfied in the MyCourse group.

In the cost-effectiveness analyses, the MyCourse group led to fewer QALYs and more reduced drinks, both at lower societal costs. Thus, MyCourse has shown to be more effective and cost-saving for number of reduced drinks.

From a societal perspective, the MyCourse group gained fewer QALYs at lower societal costs. The MyCourse intervention itself was associated with higher intervention costs than the noninteractive information brochure. Both ICERs reflected only marginally higher QALY gains in the control group. This study did not find any effect of MyCourse on QALYs. It could be hypothesized that a longer follow-up period would have been necessary for improvements in quality of life to take place in a population of cancer survivors, as their quality of life may be more directly influenced by factors pertaining to the cancer diagnosis (eg, invasiveness of cancer treatment, disease stage, cancer-related physical symptoms, and comorbidities) [42]. Therefore, we conclude that the MyCourse intervention seems more economically sustainable from a societal perspective than the noninteractive information brochure in reducing the number of drinks over a 12-month time horizon, whereas, to find evidence of possible cost-utility of a digital AM intervention, a longer follow-up period might be needed.

The difference in baseline alcohol use might be a possible explanation for the seemingly different conclusions between the incremental effect analysis and the cost-effectiveness analysis on the greater reduction of 7-day alcohol use in the MyCourse group. Even though the participants were randomized, at baseline, the participants in the MyCourse group had higher AUDIT scores and consumed more drinks on average than the participants in the control group (although the difference between the groups was not significant). At the 3-, 6-, and 12-month follow-ups, there was no significant difference in the number of drinks consumed in the past week or the AUDIT scores between the 2 groups. Thus, the larger nominal reduction in the number of drinks in the MyCourse group does not reflect a difference in the number of drinks at any of the follow-up assessments but rather a difference at baseline. This baseline difference translates differently in the incremental effect analysis, assessing differences at discrete time points, compared with the cost-effectiveness analysis, assessing differences over a period (12 months). It is also possible that, because of insufficient power, no significant difference was found in the incremental effect analysis, whereas, in the cost-effectiveness analysis, this was of lesser influence.

Wider Context

Meta-analyses of brief AM interventions [11] and AM internet interventions [10] among the general population have found that alcohol use in the past 7 days in the intervention group was reduced by approximately 20-50 g more than in the control group. Although this study did not find a significant effect on the AM rates of the MyCourse intervention over the control condition, we did find considerable reductions of approximately 70 g of ethanol at the 12-month follow-up in the control group and 120 g in the MyCourse group. These increasing reductions at longer follow-up assessments were also found in a previous digital AM intervention study [43]. In line with limited previous studies on brief and digital AM interventions [13-15], this study showed that a digital AM intervention can be cost-effective among cancer survivors. Unfortunately, because of a lack of literature, no comparisons could be made to other dedicated AM interventions for cancer survivors. A study on a telephone-counseling, combined alcohol, smoking, and depression intervention for head and neck cancer survivors also found a decrease in AUDIT scores after 6 months in both groups and no differential effect between the experimental intervention and the control group receiving only the nurse-delivered, face-to-face, 45-minute assessment, including a handout with referrals for further care (which the experimental group received as well) [44].

The lack of difference in alcohol use in the past 7 days between the MyCourse group and the control group might be due to several study aspects. It is also possible that because of the inclusion criterion of having the intention to reduce one's alcohol use, participants in both groups were highly motivated to change their alcohol use and that this obscured any effect of the MyCourse group. In addition, during this study, great efforts were made to recruit participants. At the time of recruitment in 2016-2018, AM discussion and support for cancer survivors was not well-implemented in many oncology settings; therefore, the researchers invested in informing oncology department staff on the importance and benefits of addressing AM in cancer survivors. Recruitment efforts were not only aimed at professionals; a dedicated website and a social media campaign were also in place, aiming to inform cancer survivors about the short-term benefits of AM after a cancer diagnosis while emphasizing an accepting tone to reduce possible feelings of guilt and ultimately guiding survivors to participate in the study. These recruitment efforts alone might have served as an intervention by focusing attention on AM. For cancer survivors not yet considering AM, the first step would be to address the knowledge gap on the adverse health effects of alcohol [45,46].

Second, the assessment load in this study was substantial, and the participants received multiple reminder emails and telephone calls from the researchers to fill out the survey at the respective follow-up measurement waves. Although these calls were kept as short as possible, some participants might have experienced them as part of the intervention; thus, possibly contributing to an intervention effect and making participants think regularly about their drinking behavior and feel supported [47,48]. This minimal guidance could thus have increased the AM rates in the control group. Future research should evaluate whether addressing AM in an accepting manner and with multiple repeated short reminders can encourage AM in cancer survivors.

It is possible that a true effect was not found in this study because the sample size was smaller than intended (103 instead of 114 participants). It is unlikely that the use of additional support explains the reduction of alcohol use in the control group, as very few people used any additional support. Low use of additional support was also found in a previous Dutch study on digital AM support [49]. Although the control group was as effective as the MyCourse group in reducing alcohol use in the incremental effect analyses, considering the significantly higher satisfaction rates and better economic sustainability in the MyCourse group, it would be preferable to offer the MyCourse group.

Strengths and Limitations

An important strength of this study is that the evaluation was conducted in a real-world setting; recruitment was done through both offline and web-based channels, which could plausibly be used in case of future implementation. This study succeeded in recruiting cancer survivors from a range of cancer types. The median number of times participants logged into MyCourse was high (8 times). Several sensitivity analyses have attested to the robustness of the findings. The long-term follow-up of this study showed that AM is sustained over a long period. The results should be interpreted in light of the limitations of the study. Most of this study's sample were women (86/103, 83.4%); thus, cautioning the generalization of the results to men. The participants were not blinded to their intervention allocation. Not all participants complied with the advised daily use of MyCourse for 4 weeks, and this might have influenced effects in the MyCourse group. A follow-up period of >12 months might be needed to find evidence of possible cost-utility of a digital AM intervention in cancer survivors.

Conclusions

To the best of our knowledge, this is the first study on a digital AM intervention for cancer survivors, and it showed that alcohol use was reduced by one-third in both the MyCourse and control groups and that this effect was sustained over 12 months. No significant differential effect on alcohol use between the MyCourse group and the control group was observed at the follow-ups, although cancer survivors were more satisfied in the MyCourse group. From a societal perspective, the MyCourse group seems economically more sustainable for reducing the number of drinks, as a greater reduction in the number of drinks over time was observed against lower societal costs.

Acknowledgments

The authors would like to thank Diede Kesler for her assistance in the recruitment and follow-up assessments of the participants, Yvonne Borghans for her assistance in the follow-up assessments of the participants, and professor Dr Filip Smit for his guidance

in the economic evaluations. This study was supported by grant TBOS2014–7169 from the Dutch Cancer Society (KWF Kankerbestrijding). The sponsor had no role in the design, data collection, analysis and interpretation of the data, writing of the paper, or the decision to submit for publication.

Conflicts of Interest

The intervention described in this study was developed by the Trimbos Institute (the Netherlands Institute for Mental Health and Addiction). The authors declare that they have no other competing interests.

Multimedia Appendix 1
Additional information on Methods.
[DOCX File , 16 KB - jmir_v24i2e30095_app1.docx]

Multimedia Appendix 2
Attrition and satisfaction with the intervention.
[DOCX File , 17 KB - jmir_v24i2e30095_app2.docx]

Multimedia Appendix 3
Cost-effectiveness planes and cost-effectiveness acceptability curves after Winsorization.
[DOCX File , 133 KB - jmir_v24i2e30095_app3.docx]

Multimedia Appendix 4
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1210 KB - jmir_v24i2e30095_app4.pdf]

References

1. Supramaniam R. New malignancies among cancer survivors: SEER cancer registries, 1973-2000. J Epidemiol Community Health 2008 Apr 01;62(4):375-376. [doi: 10.1136/jech.2007.063560]
2. Rock CL, Doyle C, Demark-Wahnefried W, Meyerhardt J, Courneya KS, Schwartz AL, et al. Nutrition and physical activity guidelines for cancer survivors. CA Cancer J Clin 2012;62(4):243-274 [FREE Full text] [doi: 10.3322/caac.21142] [Medline: 22539238]
3. Praud D, Rota M, Rehm J, Shield K, Zatoński W, Hashibe M, et al. Cancer incidence and mortality attributable to alcohol consumption. Int J Cancer 2016 Mar 15;138(6):1380-1387 [FREE Full text] [doi: 10.1002/ijc.29890] [Medline: 26455822]
4. Diet, nutrition, physical activity and cancer: a global perspective. World Cancer Research Fund and American Institute for Cancer Research. 2018. URL: https://www.wcrf.org/diet-and-cancer/ [accessed 2021-12-14]
5. Mowls DS, Brame LS, Martinez SA, Beebe LA. Lifestyle behaviors among US cancer survivors. J Cancer Surviv 2016 Aug;10(4):692-698. [doi: 10.1007/s11764-016-0515-x] [Medline: 26820142]
6. Shingler E, Robles LA, Perry R, Penfold C, Ness AR, Thomas S, et al. Systematic review evaluating randomized controlled trials of smoking and alcohol cessation interventions in people with head and neck cancer and oral dysplasia. Head Neck 2018 Aug;40(8):1845-1853 [FREE Full text] [doi: 10.1002/hed.25138] [Medline: 29603464]
7. Mujcic A, Blankers M, Bommelé J, Boon B, Berman AH, Verdonck-de Leeuw IM, et al. The effectiveness of distance-based interventions for smoking cessation and alcohol moderation among cancer survivors: a meta-analysis. Psychooncology 2020 Jan;29(1):49-60 [FREE Full text] [doi: 10.1002/pon.5261] [Medline: 31663182]
8. Lauridsen SV, Thomsen T, Kaldan G, Lydom LN, Tønnesen H. Smoking and alcohol cessation intervention in relation to radical cystectomy: a qualitative study of cancer patients' experiences. BMC Cancer 2017 Nov 25;17(1):793 [FREE Full text] [doi: 10.1186/s12885-017-3792-5] [Medline: 29178899]
9. Kopp LM, Gastelum Z, Guerrero CH, Howe CL, Hingorani P, Hingle M. Lifestyle behavior interventions delivered using technology in childhood, adolescent, and young adult cancer survivors: a systematic review. Pediatr Blood Cancer 2017 Dec;64(1):13-17. [doi: 10.1002/pbc.26166] [Medline: 27468131]
10. Riper H, Hoogendoorn A, Cuijpers P, Karyotaki E, Boumparis N, Mira A, et al. Effectiveness and treatment moderators of internet interventions for adult problem drinking: an individual patient data meta-analysis of 19 randomised controlled trials. PLoS Med 2018 Dec 18;15(12):e1002714 [FREE Full text] [doi: 10.1371/journal.pmed.1002714] [Medline: 30562347]
11. Kaner E, Beyer F, Dickinson HO, Pienaar E, Campbell F, Schlesinger C, et al. Effectiveness of brief alcohol interventions in primary care populations. Cochrane Database Syst Rev 2007 Apr 18;2(2):CD004148. [doi: 10.1002/14651858.CD004148.pub3] [Medline: 17443541]
12. Kaner EF, Beyer FR, Garnett C, Crane D, Brown J, Muirhead C, et al. Personalised digital interventions for reducing hazardous and harmful alcohol consumption in community-dwelling populations. Cochrane Database Syst Rev 2017 Dec 25;9:CD011479. [doi: 10.1002/14651858.CD011479.pub2] [Medline: 28944453]

13. Angus C, Latimer N, Preston L, Li J, Purshouse R. What are the implications for policy makers? A systematic review of the cost-effectiveness of screening and brief interventions for alcohol misuse in primary care. Front Psychiatry 2014 Sep 01;5:114 [FREE Full text] [doi: 10.3389/fpsyt.2014.00114] [Medline: 25225487]
14. Angus C, Li J, Romero-Rodriguez E, Anderson P, Parrott S, Brennan A. Cost-effectiveness of strategies to improve delivery of brief interventions for heavy drinking in primary care: results from the ODHIN trial. Eur J Public Health 2019 Apr 01;29(2):219-225. [doi: 10.1093/eurpub/cky181] [Medline: 30239676]
15. Drost RM, Paulus AT, Jander AF, Mercken L, de Vries H, Ruwaard D, et al. A web-based computer-tailored alcohol prevention program for adolescents: cost-effectiveness and intersectoral costs and benefits. J Med Internet Res 2016 Apr 21;18(4):e93 [FREE Full text] [doi: 10.2196/jmir.5223] [Medline: 27103154]
16. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating digital health interventions: key questions and approaches. Am J Prev Med 2016 Nov;51(5):843-851 [FREE Full text] [doi: 10.1016/j.amepre.2016.06.008] [Medline: 27745684]
17. Ross J, Stevenson F, Lau R, Murray E. Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). Implement Sci 2016 Oct 26;11(1):146 [FREE Full text] [doi: 10.1186/s13012-016-0510-7] [Medline: 27782832]
18. Krebber AM, Buffart LM, Kleijn G, Riepma IC, de Bree R, Leemans CR, et al. Prevalence of depression in cancer patients: a meta-analysis of diagnostic interviews and self-report instruments. Psychooncology 2014 Feb 16;23(2):121-130 [FREE Full text] [doi: 10.1002/pon.3409] [Medline: 24105788]
19. Stanton AL. Psychosocial concerns and interventions for cancer survivors. J Clin Oncol 2006 Nov 10;24(32):5132-5137. [doi: 10.1200/JCO.2006.06.8775] [Medline: 17093275]
20. Mujcic A, Blankers M, Boon B, Engels R, van Laar M. Internet-based self-help smoking cessation and alcohol moderation interventions for cancer survivors: a study protocol of two RCTs. BMC Cancer 2018 Apr 02;18(1):364 [FREE Full text] [doi: 10.1186/s12885-018-4206-z] [Medline: 29609554]
21. Riper H, Kramer J, Smit F, Conijn B, Schippers G, Cuijpers P. Web-based self-help for problem drinkers: a pragmatic randomized trial. Addiction 2008 Feb;103(2):218-227. [doi: 10.1111/j.1360-0443.2007.02063.x] [Medline: 18199300]
22. Atkins DC, Baldwin SA, Zheng C, Gallop RJ, Neighbors C. A tutorial on count regression and zero-altered count models for longitudinal substance use data. Psychol Addict Behav 2013 Mar;27(1):166-177 [FREE Full text] [doi: 10.1037/a0029508] [Medline: 22905895]
23. Sobell L, Sobell M. Timeline follow-back. In: Measuring Alcohol Consumption. Totowa, NJ: Humana Press; 1992:41-72.
24. Heatherton T, Kozlowski L, Frecker R, Fagerström KO. The Fagerström Test for nicotine dependence: a revision of the Fagerström Tolerance Questionnaire. Br J Addict 1991 Sep;86(9):1119-1127. [doi: 10.1111/j.1360-0443.1991.tb01879.x] [Medline: 1932883]
25. Crowne DP, Marlowe D. A new scale of social desirability independent of psychopathology. J Consult Psychol 1960;24(4):349-354. [doi: 10.1037/h0047358] [Medline: 13813058]
26. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med 2001 Jul;33(5):337-343. [doi: 10.3109/07853890109002087] [Medline: 11491192]
27. Ware JJ, Sherbourne C. The MOS 36-ltem Short-Form Health Survey (SF-36). Medical Care 1992;30(6):473-483. [doi: 10.1097/00005650-199206000-00002]
28. Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. Med Care 2004 Sep;42(9):851-859. [doi: 10.1097/01.mlr.0000135827.18610.0d] [Medline: 15319610]
29. Kriz D, Nübling R, Steffanowski A, Wittmann W, Schmidt J. Patientenzufriedenheit in der stationären Rehabilitation: Psychometrische Reanalyse des ZUF-8 auf der Basis multizentrischer Stichproben verschiedener Indikation. Zeitschrift für Medizinische Psychol. 2008. URL: http://www.redi-bw.de/db/ebsco.php/search.ebscohost.com/login.aspx?direct=true&db=pdx&AN=0210616&site=ehost-live [accessed 2015-10-17]
30. Riper H, Blankers M, Hadiwijaya H, Cunningham J, Clarke S, Wiers R, et al. Effectiveness of guided and unguided low-intensity internet interventions for adult alcohol misuse: a meta-analysis. PLoS One 2014;9(6):e99912 [FREE Full text] [doi: 10.1371/journal.pone.0099912] [Medline: 24937483]
31. Van Dieren S, Eskes A. Economische evaluaties in de gezondheidszorg. TVZ 2018 Dec 18;128(6):56-57. [doi: 10.1007/s41184-018-0215-2]
32. Medicijnkosten. Zorginstituut Nederland. 2019. URL: https://www.medicijnkosten.nl/ [accessed 2019-11-01]
33. van Roijen L, van Straten A, Tiemens B, Donker M. Manual Trimbos/iMTA Questionnaire for costs associated with psychiatric illness (TIC-P) (in Dutch). ResearchGate. 2002. URL: https://www.researchgate.net/publication/254758194_Manual_TrimbosiMTA_Questionnaire_for_Costs_Associated_with_Psychiatric_Illness_TIC-P_in_Dutch [accessed 2021-12-13]
34. Blankers M, Koeter MW, Schippers GM. Internet therapy versus internet self-help versus no treatment for problematic alcohol use: a randomized controlled trial. J Consult Clin Psychol 2011 Jun;79(3):330-341. [doi: 10.1037/a0023498] [Medline: 21534652]
35. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2018.

36. Buuren SV, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. J Stat Soft 2011;45(3):1-68. [doi: 10.18637/jss.v045.i03]
37. Koller M. robustlmm: an R package for robust estimation of linear mixed-effects models. J Stat Soft 2016;75(6). [doi: 10.18637/jss.v075.i06]
38. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Soft 2015;67(1). [doi: 10.18637/jss.v067.i01]
39. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, CHEERS Task Force. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. Cost Eff Resour Alloc 2013 Mar 25;11(1):6 [FREE Full text] [doi: 10.1186/1478-7547-11-6] [Medline: 23531194]
40. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. Oxford, United Kingdom: Oxford university press; 2015.
41. Lamers LM, Stalmeier PF, McDonnell J, Krabbe PF, van Busschbach JJ. [Measuring the quality of life in economic evaluations: the Dutch EQ-5D tariff]. Ned Tijdschr Geneeskd 2005 Jul 09;149(28):1574-1578. [Medline: 16038162]
42. Götze H, Taubenheim S, Dietz A, Lordick F, Mehnert A. Comorbid conditions and health-related quality of life in long-term cancer survivors-associations with demographic and medical characteristics. J Cancer Surviv 2018 Oct 10;12(5):712-720. [doi: 10.1007/s11764-018-0708-6] [Medline: 30097854]
43. Freyer-Adam J, Baumann S, Haberecht K, Tobschall S, Bischof G, John U, et al. In-person alcohol counseling versus computer-generated feedback: results from a randomized controlled trial. Health Psychol 2018 Dec;37(1):70-80. [doi: 10.1037/hea0000556] [Medline: 28967769]
44. Duffy SA, Ronis DL, Valenstein M, Lambert MT, Fowler KE, Gregory L, et al. A tailored smoking, alcohol, and depression intervention for head and neck cancer patients. Cancer Epidemiol Biomarkers Prev 2006 Nov;15(11):2203-2208 [FREE Full text] [doi: 10.1158/1055-9965.EPI-05-0880] [Medline: 17119047]
45. Coomber K, Mayshak R, Curtis A, Miller PG. Awareness and correlates of short-term and long-term consequences of alcohol use among Australian drinkers. Aust N Z J Public Health 2017 Jun 22;41(3):237-242. [doi: 10.1111/1753-6405.12634] [Medline: 28110495]
46. LoConte NK, Brewster AM, Kaur JS, Merrill JK, Alberg AJ. Alcohol and cancer: a statement of the American Society of Clinical Oncology. J Clin Oncol 2018 Jan 01;36(1):83-93. [doi: 10.1200/JCO.2017.76.1155] [Medline: 29112463]
47. Meier E, Miller MB, Lombardi N, Leffingwell T. Assessment reactivity: a randomized controlled trial of alcohol-specific measures on alcohol-related behaviors. Addict Behav 2017 Apr;67:44-48 [FREE Full text] [doi: 10.1016/j.addbeh.2016.11.025] [Medline: 27992833]
48. Miles LM, Rodrigues AM, Sniehotta FF, French DP. Asking questions changes health-related behavior: an updated systematic review and meta-analysis. J Clin Epidemiol 2020 Jul;123:59-68 [FREE Full text] [doi: 10.1016/j.jclinepi.2020.03.014] [Medline: 32229251]
49. Riper H, Kramer J, Conijn B, Smit F, Schippers G, Cuijpers P. Translating effective web-based self-help for problem drinking into the real world. Alcohol Clin Exp Res 2009 Aug;33(8):1401-1408. [doi: 10.1111/j.1530-0277.2009.00970.x] [Medline: 19413646]

Abbreviations

AM: alcohol moderation
AUDIT: Alcohol Use Disorders Identification Test
CEAC: cost-effectiveness acceptability curve
EQ-5D-5L: 5-level EuroQol 5 Dimension
ICER: incremental cost-effectiveness ratio
MCSDS: Marlowe-Crowne Social Desirability Scale
QALY: quality-adjusted life year
RCT: randomized controlled trial
SF-6D: Short Form 6-dimension
TLFB: Timeline Followback
ZUF-8: Patient Satisfaction Questionnaire

Edited by R Kukafka; submitted 05.05.21; peer-reviewed by K Matsumoto, K Champion; comments to author 02.09.21; revised version received 05.11.21; accepted 15.11.21; published 01.02.22.

Please cite as:
Mujcic A, Blankers M, Boon B, Berman AH, Riper H, van Laar M, Engels R
Effectiveness, Cost-effectiveness, and Cost-Utility of a Digital Alcohol Moderation Intervention for Cancer Survivors: Health Economic Evaluation and Outcomes of a Pragmatic Randomized Controlled Trial
J Med Internet Res 2022;24(2):e30095
URL: https://www.jmir.org/2022/2/e30095
doi:10.2196/30095
PMID:35103605

Original Paper

Implementation of an Electronic Patient-Reported Outcome App for Health-Related Quality of Life in Breast Cancer Patients: Evaluation and Acceptability Analysis in a Two-Center Prospective Trial

Joachim Graf[1*], DPhil; Nina Sickenberger[2*], MD; Katharina Brusniak[2], MD; Lina Maria Matthies[3], MD; Thomas M Deutsch[3], MD; Elisabeth Simoes[4,5], Prof Dr; Claudia Plappert[1], Prof Dr; Lucia Keilmann[6], MD; Andreas Hartkopf[4], Prof Dr; Christina Barbara Walter[4], MD; Markus Hahn[4], Prof Dr; Tobias Engler[4], MD; Stephanie Wallwiener[2], Prof Dr; Florian Schuetz[3,7], Prof Dr; Peter A Fasching[8,9], Prof Dr; Andreas Schneeweiss[3], Prof Dr; Sara Yvonne Brucker[4,5], Prof Dr; Markus Wallwiener[2,3], Prof Dr

[1]Institute for Health Sciences, Section of Midwifery Science, University Hospital Tübingen, Tübingen, Germany

[2]Hospital for General Obstetrics and Gynecology, University Hospital Heidelberg, Heidelberg, Germany

[3]National Center for Tumor Diseases, University Hospital and German Cancer Research Center, Heidelberg, Germany

[4]Department of Women's Health, University Hospital Tübingen, Tübingen, Germany

[5]Department of Women's Health, Research Institute for Women's Health, University Hospital Tübingen, Tübingen, Germany

[6]Department of Obstetrics and Gynecology, University Hospital Ludwig-Maximilians-University Munich, München, Germany

[7]Diakonissen-Stiftungs-Krankenhaus Speyer, Speyer, Germany

[8]Department of Gynecology and Obstetrics, University Hospital Erlangen, Erlangen, Germany

[9]University Breast Center Franconia, Comprehensive Cancer Center Erlangen-EMN, Friedrich-Alexander University Erlangen-Nuremberg, Erlangen, Germany

[*]these authors contributed equally

Corresponding Author:
Markus Wallwiener, Prof Dr
Hospital for General Obstetrics and Gynecology
University Hospital Heidelberg
Im Neuenheimer Feld 440
Heidelberg, 69120
Germany
Phone: 49 06221 ext 56
Email: markus.wallwiener@gmail.com

Abstract

Background: One in eight women is diagnosed with breast cancer in the course of their life. As systematic palliative treatment has only a limited effect on survival rates, the concept of health-related quality of life (HRQoL) was developed for measurement of patient-centered outcomes. Various studies have already demonstrated the reliability of paper-based patient-reported outcome (pPRO) and electronic patient-reported outcome (ePRO) surveys and that the 2 means of assessment are equally valid.

Objective: The aim of this study was to analyze the acceptance and evaluation of a tablet-based ePRO app for breast cancer patients and to examine its suitability, effort, and difficulty in the context of HRQoL and sociodemographic factors.

Methods: Overall, 106 women with adjuvant or advanced breast cancer were included in a 2-center study at 2 major university hospitals in Germany. Patients were asked to answer HRQoL and PRO questionnaires both on a tablet on-site using a specific eHealth assessment website and on paper. The suitability, effort, and difficulty of the app and self-reported technical skills were also assessed. Only the results of the electronically acquired data are presented here. The results of the reliability of the pPRO data have already been published elsewhere.

Results: Patients regarded the ePRO assessment as more suitable (80/106, 75.5%), less stressful (73/106, 68.9%), and less difficult (69/106, 65.1%) than pPRO. The majority of patients stated that ePRO assessment improves health care in hospitals (87/106, 82.1%). However, evaluation of ePROs depended on the level of education (P=.003) in the dimensions of effort and

difficulty (regression analysis). The app was rated highly in all categories. HRQoL data and therapy setting did not show significant correlations with the app's evaluation parameters.

Conclusions: The results indicate that ePRO surveys are feasible for measuring HRQoL in breast cancer patients and that those patients prefer ePRO assessment to pPRO assessment. It can also be seen that patients consider ePRO assessment to improve hospital health care. However, studies with larger numbers of patients are needed to develop apps that address the needs of patients with lower levels of education and technical skills.

(J Med Internet Res 2022;24(2):e16128) doi:10.2196/16128

KEYWORDS

eHealth; electronic patient-reported outcomes; evaluation; acceptability; breast cancer

Introduction

Breast Cancer: Epidemiological Relevance

With about 70,000 new cases each year, breast cancer is the most common cancer in Germany. One in eight women is diagnosed with breast cancer in the course of their life [1]. Due to great advances in cancer therapy options, the relative 5-year survival rate after initial diagnosis has increased to 88% [2]. While patients in adjuvant situations have an improved prognosis, patients with metastatic disease remain incurable and are hence treated with palliative care. Since currently systemic palliative treatment has a limited effect on survival rates, the concept of health-related quality of life (HRQoL) and measurement of patient-reported outcomes (PROs) are gaining increasing importance in the therapy of progressive diseases, such as breast cancer, especially in the metastatic setting [3-7]. Under this directive, the issue of how HRQoL data can be collected as efficiently and accurately as possible in real-world settings is gaining importance [8].

PROs as a Holistic Addition to Clinic-Reported Outcomes

Drug evaluation studies have focused on clinical endpoints (clinic-reported outcomes), such as overall survival and progression-free survival, for years. Yet, PROs are becoming increasingly important to verify and compare the efficacy of different chemotherapeutic interventions in drug evaluation studies, not least owing to legal regulations [9]. This fact is confirmed by the enormous increase in studies publishing PRO data over the last few decades [10]. A PRO is widely defined as "any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else" [11]. The concept of PRO takes into account the patient's point of view concerning health status, therapy intervention, negative side effects, mental and functional components, satisfaction with care, drug adherence, and impact of progressive disease [9-14].

Potential of Electronic PROs

Strong willingness to use technology within the population and existing infrastructure represents the rationale for digitalization in many sectors, including health [15]. While PROs are still routinely captured via paper-based methods, technical progress is gradually allowing more PRO data to be collected in the form of electronic PROs (ePROs), that is, via tablet computers or smartphones [16]. To ensure that patients are able to deal with ePRO data capture, studies that prove reliability and acceptance are needed. Various studies have already demonstrated that both paper-based PRO (pPRO) and ePRO surveys are reliable and equally valid means of assessment [4,17-21]. Nevertheless, knowledge about the detailed evaluation of ePRO apps and information regarding patient acceptance, feasibility, and barriers are still limited [22], especially in relation to sociodemographic aspects, health status, and technical skills [7,23-26]. It also remains unclear how ePRO questionnaires are accepted and evaluated in breast cancer patients, in whom high patient satisfaction with use and usability are important implementation prerequisites for capturing real-world evidence in routine clinical care.

Aims and Objectives

The aim of this study was to analyze the acceptance and evaluation of a tablet-based ePRO app for breast cancer patients. More specifically, we investigated how suitable patients maintain an app that is used to collect HRQoL data, whether they find it diffuse or difficult to use, and how to evaluate individual aspects of the questionnaire. To determine whether the HRQoL survey app can be used in all breast cancer patients, we also examined whether the app's suitability, effort, and difficulty ratings were dependent on HRQoL and sociodemographic factors (age, educational status, and computer skills).

Methods

Study Design and Sample

The methodology has already been described in detail elsewhere [19,21]. Here, it was shown that HRQoL can be validly assessed by the related tool, since no significant differences in response behavior between pPROs and ePROs were found with regard to reliability in both the European Organization for the Research and Treatment of Cancer core quality of life questionnaire (EORTC QLQ-C30) and the Functional Assessment of Cancer Therapy-Breast Cancer (FACT-B) questionnaires. In this study, the feasibility and acceptability will be evaluated. For digital assessment, we used a web-based solution PiiA (patient interactively informs doctor), allowing patients to answer the HRQoL assessment on a tablet after receiving anonymized user credentials [19,21]. Patients were recruited as a part of the ePROCOM (electronic Patient-Reported Outcomes and Compliance Analysis) and the PEPPER (Patient Engagement Breast Cancer) study. While the ePROCOM study aims to evaluate general patient acceptance and practicability of a

web-based app for a PRO questionnaire for patients with adjuvant or metastatic breast cancer, the PEPPER study aims to evaluate the impact of web-based PROs and pPROs for health care services. The inclusion criteria were female gender, full legal age, proven diagnosis of breast cancer in an adjuvant or metastatic setting, sufficient language skills in German, and signed declaration of consent. The exclusion criterion was participation in other studies to minimize the burden of questionnaires. Patients were asked to complete the questionnaire during an outpatient visit at the hospital under the supervision of an attending physician. From July 2015 to May 2016, questionnaires were completed by a total of 106 female adjuvant and metastatic breast cancer patients treated consecutively at the Department of Women's Health in Tubingen, Germany, and the National Center for Tumor Diseases (NCT) in Heidelberg, Germany. The study was designed as a 2-center prospective trial (Tübingen and Heidelberg). Ethical approval was granted by the Ethics Committee of the University of Heidelberg (S-569/2015) and the Ethics Committee of the University of Tübingen (project number 089/2015B01).

Questionnaires

All patients were required to complete both the ePRO and pPRO versions of the EORTC QLQ-C30 and FACT-B HRQoL questionnaires. The reliability of a tablet-based ePRO app of both questionnaires has already been analyzed (the results of the reliability analysis have been published previously [19,21]). Furthermore, patients were questioned about pre-existing technical skills, their willingness to use ePROs, potential barriers in relation to their health status, and socioeconomic variables [7,26]. Thereafter, patients were asked to evaluate the app and its handling, which is reported in the current paper. The questionnaires for measuring socioeconomic status and evaluating the app were developed by our own research group. Regarding the evaluation of the app, patients assessed the ePRO questionnaires in terms of suitability, effort, and difficulty compared with the pPRO questionnaires. They also mentioned whether the introduction of ePROs was thought to have a positive impact on the quality of health care and how they considered the app in terms of usability, graphic design, and applicability. Patients were informed about the aims of the study and were asked for their consent ex ante.

Statistical Analyses

All statistical analyses were conducted using IBM SPSS Statistics (Version 24). First, a frequency analysis was performed to determine the descriptive sociodemographic characteristics of the patients. Subsequently, the mean values and dispersion parameters of the variable age were calculated, and then, the frequencies of the individual dimensions of the variable educational attainment were determined. Thereafter, HRQoL from the EORTC QLQ-C30 questionnaire [21] and computer skills of the patients were assessed before the evaluation sheets were analyzed descriptively. Thereby, stratification was done by treatment setting, using the Mann-Whitney *U* test to test the significance of the identified frequency differences. Subsequently, multivariable regression analyses were conducted on the 3 aforementioned target outcomes. The aim of regression analysis was to determine whether the regression models showed statistically significant relationships between the evaluation dimensions of suitability, effort, and difficulty and between socioeconomic variables (age and educational attainment), computer skills, the treatment regimen (metastatic vs adjuvant), and HRQoL. For each level of ordinal variables, dummy variables were created in multivariable regression analyses. In all analyses, *P* values <.05 (2-tailed) were considered indicative of statistically significant differences (α=.05).

Results

Sociodemographic Variables

Table 1 shows the sociodemographic characteristics of the study group stratified by therapy setting with 76 (72%) patients in adjuvant therapy and 30 (28%) with metastatic disease, as well as their HRQoL, their self-related computer skills, and their computer use experience in years. There were no significant differences between adjuvant and metastatic patients. The mean age was 49.4 years in the adjuvant group and 53.9 years in the metastatic group. Nearly half of the patients (adjuvant patients: 38/76, 50%; metastatic patients: 14/30, 47%) had a higher level of education (advanced technical graduation or high school diploma). The mean HRQoL score was approximately 60 points in both groups (where 0 represents the worst value and 100 the highest value). Among all patients, more than three-quarters rated their computer skills as advanced or professional (adjuvant patients: 54/76, 71%; metastatic patients: 20/30, 67%), while the mean time of computer use was more than 10 years in both groups.

Table 1. Sociodemographic characteristics of the patients.

Sociodemographic variables	Adjuvant therapy group (n=76, 72%)	Metastatic situation group (n=30, 28%)	*P* value
Age (years)			.18
Mean value (SD)	49.39 (10.28)	53.93 (13.94)	
Median value (minimum-maximum)	50.0 (27-73)	52.0 (33-84)	
Level of education, n (%)			.37
No qualification	0 (0)	1 (3)	
Main/secondary school graduation	30 (39)	11 (37)	
Advanced technical graduation	11 (14)	8 (27)	
High school diploma (“Abitur”)	27 (36)	6 (20)	
Not specified	8 (11)	4 (13)	
HRQoL[a] (overall HRQoL from EORTC QLQ-C30[b])			.16
Mean value (SD)	63.51 (23.26)	57.77 (19.27)	
Median value (minimum-maximum)	67 (17-100)	62.5 (17-100)	
Computer skills (self-perception by the patients)			.61
Mean value (SD)	2.82 (0.64)	2.69 (0.62)	
Computer skills level, n (%)			
Beginner (1)	2 (3)	2 (7)	
Basic (2)	16 (21)	4 (13)	
Advanced (3)	47 (62)	20 (67)	
Professional (4)	7 (9)	0 (0)	
Not specified	4 (5)	4 (13)	
Computer use (years)			.18
Mean value (SD)	17.50 (7.09)	14.22 (9.59)	
Median value (minimum-maximum)	18 (0-35)	15 (0-35)	

[a]HRQoL: health-related quality of life.

[b]EORTC QLQ-C30: European Organization for the Research and Treatment of Cancer core quality of life questionnaire.

Evaluation of Suitability, Effort, and Difficulty: Comparison Between ePRO and pPRO Surveys

Figure 1 shows how the patients evaluated the app in comparison to the paper version. We examined how the platform was rated in terms of suitability compared to the pPRO survey on a 5-point Likert scale. A rating of 3 was considered as comparable suitability between the 2 assessment strategies, whereas a rating of 1 or 2 was regarded as low suitability and a rating of 4 or 5 was regarded as high suitability. Three-quarters of the patients (80/106, 75.5%) reported that the ePRO survey on the EORTC QLQ-C30 and the FACT-B questionnaire was more appropriate than the pPRO survey.

Similar results were obtained in the dimensions effort and difficulty. One-quarter of the patients stated that completing the ePRO sheets was as stressful (27/106, 25.5%) and as difficult (29/106, 27.4%) as completing the HRQoL questionnaires on paper, whereas 68.9% (73/106) of the patients rated the ePRO survey as less stressful and 65.1% (69/106) as less difficult than the pPRO survey. The proportion of patients who rated the ePRO survey worse than the pPRO survey was negligible (Figure 2).

Overall, 82.1% (87/106) of patients said that the introduction of the ePRO survey improved health care in hospitals, 16.0% (17/106) of patients said that the introduction of the ePRO survey was associated with deterioration in health care, and 1.9% (2/106) of patients said that the introduction of the ePRO survey had no impact on health care in hospitals.

Figure 1. Suitability of the electronic patient-reported outcome survey in relation to the paper-based patient-reported outcome (pPRO) survey.

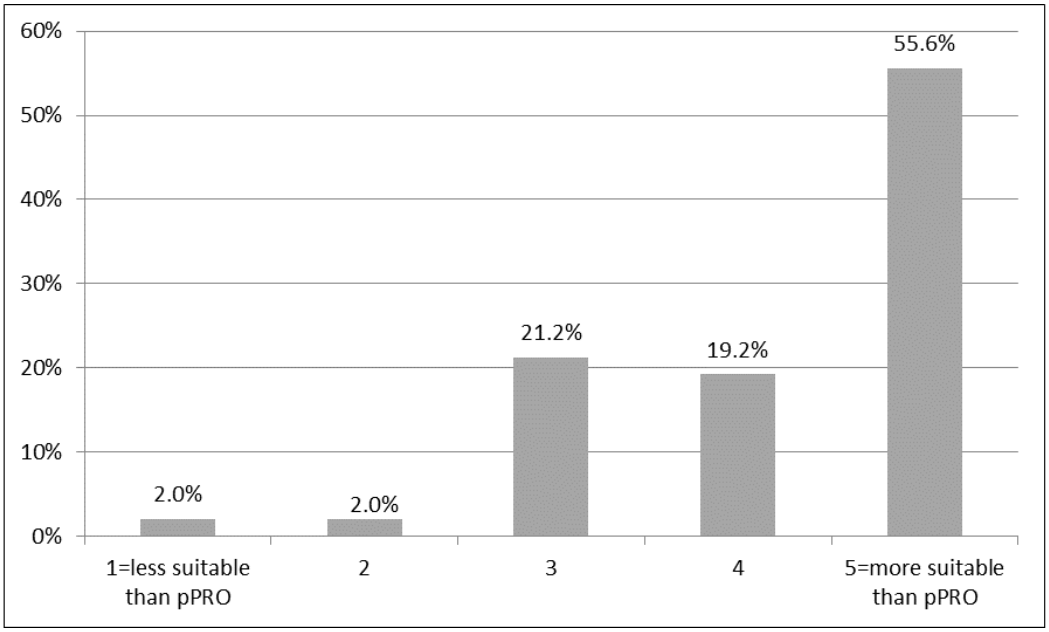
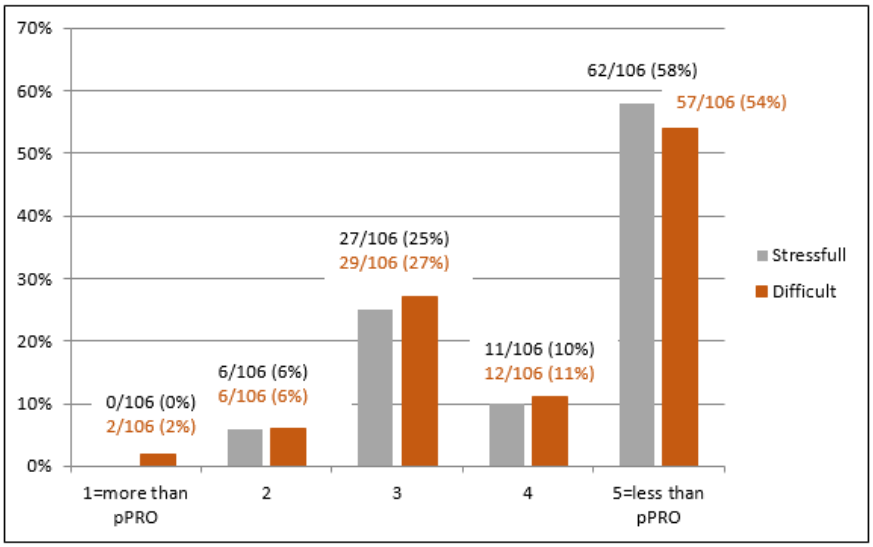


Figure 2. Effort and difficulty of the electronic patient-reported outcome survey in relation to the paper-based patient-reported outcome (pPRO) survey.



Influential Factors in ePRO Evaluation

For the suitability, effort, and difficulty of the ePRO survey, we examined whether the respective evaluation was influenced by socioeconomic factors, HRQoL, therapeutic setting, self-assessed computer knowledge, or experience in using computer technology. In the suitability dimension, no statistically significant correlations were found in the multivariable regression analysis. Statistically significant regression correlations were found between evaluation in the effort and difficulty dimensions and the educational level, as well as the time span of computer technology use. With higher levels of education and increasing time of using computer technology, completing the ePRO survey was more often reported as requiring less effort and being less difficult than completing paper questionnaires. By contrast, age, HRQoL, therapy setting (metastasized vs adjuvant therapy), and computer skills did not influence the response behavior in the evaluation. A total of 10.5% of the assessments could be attributed to the level of education in the effort dimension assessment, and a total of 14.5% in the dimension of difficulty, while time span of computer use only influenced the evaluation with 0.2% in both dimensions (Table 2).

Table 2. Multivariable regression analyses on suitability, effort, and difficulty of the electronic patient-reported outcome assessment app.

Variable	*R*	R^2	*P* value	95% CI
Dependent variable: Suitability of the ePRO[a] assessment				
Therapy setting	0.152	0.0231	.55	−0.356 to 0.659
Age	0.007	<0.0001	.51	−0.014 to 0.029
Education	−0.132	0.0174	.24	−0.353 to 0.090
Time of computer use	0.034	0.0012	.06	−0.001 to 0.069
Computer skills	−0.098	0.0096	.64	−0.519 to 0.323
HRQoL[b]	0.005	<0.0001	.97	−0.010 to 0.010
Dependent variable: Effort of the ePRO assessment app				
Therapy setting	0.136	0.0185	.58	−0.353 to 0.624
Age	0.014	0.0002	.16	−0.006 to 0.034
Education	−0.324	0.1050	.003[c]	−0.536 to −0.113
Time of computer use	0.043	0.0018	.01[c]	0.009 to 0.076
Computer skills	0.004	<0.0001	.98	−0.396 to 0.404
HRQoL	0.008	<0.0001	.12	−0.002 to 0.017
Dependent variable: Difficulty of the ePRO assessment app				
Therapy setting	0.071	0.0050	.80	−0.492 to 0.633
Age	0.006	<0.0001	.62	−0.017 to 0.028
Education	−0.381	0.1451	.003[c]	−0.626 to −0.136
Time of computer use	0.040	0.0016	.04[c]	0.002 to 0.079
Computer skills	0.093	0.0086	.69	−0.365 to 0.551
HRQoL	0.004	<0.0001	.50	−0.007 to 0.015

[a]ePRO: electronic patient-reported outcome.

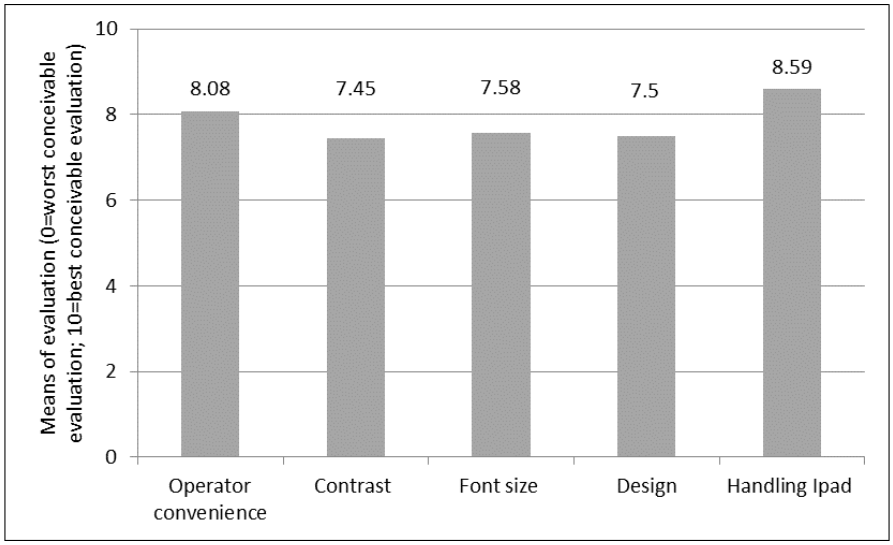
[b]HRQoL: health-related quality of life.

[c]Statistically significant difference.

Evaluation of the App's Usability

Figure 3 shows the mean values of the usability evaluation. All 5 dimensions had high to very high ratings. The dimensions operator convenience, contrast, font size, and design were scored between 7.4 and 8.1, while handling was scored at 8.6 on average.

Figure 3. Application usability in different aspects of user experience.



Discussion

Principal Findings

The results indicated that ePRO surveys are also applicable for measuring HRQoL in breast cancer patients with metastatic disease or under adjuvant therapy. The app was rated as more suitable, as requiring an equal or lesser degree of effort, and as being equally or less difficult than the pPRO survey. The evaluation of usability and applicability showed high to very high ratings. Thus, it could be shown that the ePRO survey can be used even in patients with a high burden of disease as well as in older patients, as HRQoL and age did not affect the evaluation. This is a significant finding, which extends the previous state of research, as it previously appeared unclear whether there were barriers to using ePROs in elderly and metastatic patients [7]. The influence of educational level on the evaluation was significant but rather small.

Comparison With Prior Work

Although ePRO apps are being adopted more frequently, paper-based surveys of PROs still predominate in clinical research because reliable electronically validated questionnaires are lacking. This is why knowledge about the detailed evaluation of ePRO apps and information regarding patient acceptance, feasibility, and barriers are still limited [22], although the potential of ePROs is high and the experience is very satisfactory so far [27]. The results of this study basically confirm the results of few existing studies [4,24,28,29]. Wintner et al likewise showed that cancer patients preferred ePRO questionnaires over pPRO questionnaires [28]. A high rating could also be found for electronic psycho-oncological screening instruments in breast cancer patients; here, the acceptance was greater than that of the paper-pencil screening [24]. However, only results from very small patient populations are available for breast cancer patients with metastatic disease or under adjuvant therapy. Both a Japanese and a German research group were able to demonstrate positive effects in this area. However, the number of patients included was less than 20 [4,29]. The current results therefore represent a unique characteristic, as we were able to demonstrate for the first time that ePRO surveys are also well received and better evaluated than paper-based surveys by patients with a high burden of disease in a larger collective. Other studies have not yet focused on the factors that influence app evaluation (and thus the response behavior of patients [22]). The fact that the level of education and the time span of using computer technology influence the evaluation of ePROs confirms the findings of our team, as the willingness to use such an app is also influenced by socioeconomic factors and computer skills [7,26].

Limitations

Despite positive results, some limitations of the study design and methodological implementation should be mentioned, which could possibly reduce the representativeness of the data. Patients were required to complete questionnaires during an outpatient hospital visit. The phenomenon of socially desirable response behavior might have influenced the evaluation results, such that the app might have been rated differently by patients in their home environment. It also needs to be noted that there might have been selection bias, as we did not examine whether the HRQoL was lower and the psychological distress was higher in those patients who could not be motivated to participate in the study. Only patients who were already technically inclined might have been willing to participate. Therefore, it remains unclear how acceptance differs from that in patients who only display a low level of use willingness [7,26].

Conclusions

Although digital assessment of HRQoL is constantly being adopted in clinical research and clinical routine, knowledge about the evaluation and acceptance of ePRO apps in breast cancer patients with a high burden of disease is insufficient. The results of this study indicate that breast cancer patients with metastatic disease and those under adjuvant therapy prefer ePRO surveys to pPRO surveys. However, the evaluation of ePROs depends on the level of education and the patient's computer skills and experience. Here, studies with larger collectives are needed to develop low-threshold offers that make ePRO surveys usable for all patient groups in both clinical and home settings and to better understand the needs of patients with a higher disease burden.

Acknowledgments

This study was supported by a grant of the Klaus Tschira Stiftung.

Conflicts of Interest

SYB has received honoraria from Roche, Novartis, Astraseneca, Sanofi, Medtronic, and Storz. CBW has received honoraria from Roche and Novartis. AS reports grants from Celgene, Roche, and AbbVie, as well as personal fees from Roche, AstraZeneca, Celgene, Pfizer, Novartis, MSD, Tesaro and Lilly.

References

1. Krebs in Deutschland 2011/2012. Krebsdaten. 2015. URL: https://www.krebsdaten.de/Krebs/DE/Content/Publikationen/Krebs_in_Deutschland/kid_2015/krebs_in_deutschland_2015.pdf?__blob=publicationFile [accessed 2022-01-02]
2. Brustkrebs (Mammakarzinom) ICD-10 C50. Krebsdaten. URL: http://www.krebsdaten.de/Krebs/DE/Content/Krebsarten/Brustkrebs/brustkrebs_node.htm [accessed 2022-01-02]
3. Cortés J, Baselga J, Im Y, Im S, Pivot X, Ross G, et al. Health-related quality-of-life assessment in CLEOPATRA, a phase III study combining pertuzumab with trastuzumab and docetaxel in metastatic breast cancer. Ann Oncol 2013 Oct;24(10):2630-2635 [FREE Full text] [doi: 10.1093/annonc/mdt274] [Medline: 23868905]
4. Wallwiener M, Heindl F, Brucker SY, Taran F, Hartkopf A, Overkamp F, et al. Implementation and Feasibility of Electronic Patient-Reported Outcome (ePRO) Data Entry in the PRAEGNANT Real-Time Advanced and Metastatic Breast Cancer Registry. Geburtshilfe Frauenheilkd 2017 Aug;77(8):870-878 [FREE Full text] [doi: 10.1055/s-0043-116223] [Medline: 28845051]
5. Karamouzis MV, Ioannidis G, Rigatos G. Quality of life in metastatic breast cancer patients under chemotherapy or supportive care: a single-institution comparative study. Eur J Cancer Care (Engl) 2007 Sep;16(5):433-438. [doi: 10.1111/j.1365-2354.2006.00771.x] [Medline: 17760930]
6. Bitton A, Onega T, Tosteson ANA, Haas JS. Toward a better understanding of patient-reported outcomes in clinical practice. Am J Manag Care 2014 Apr;20(4):281-283 [FREE Full text] [Medline: 24884859]
7. Graf J, Simoes E, Wißlicen K, Rava L, Walter CB, Hartkopf A, et al. Willingness of Patients with Breast Cancer in the Adjuvant and Metastatic Setting to Use Electronic Surveys (ePRO) Depends on Sociodemographic Factors, Health-related Quality of Life, Disease Status and Computer Skills. Geburtshilfe Frauenheilkd 2016 May;76(5):535-541 [FREE Full text] [doi: 10.1055/s-0042-105872] [Medline: 27239062]
8. Wallwiener M, Simoes E, Sokolov AN, Brucker SY, Fasching PA, Graf J. Health-related Quality of Life in Metastatic and Adjuvant Breast Cancer Patients. Geburtshilfe Frauenheilkd 2016 Oct;76(10):1065-1073 [FREE Full text] [doi: 10.1055/s-0042-113188] [Medline: 27761027]
9. Klakow-Franck R. [The relevance of quality of life for the work of the Federal Joint Committee]. Z Evid Fortbild Qual Gesundhwes 2014;108(2-3):151-156. [doi: 10.1016/j.zefq.2014.02.007] [Medline: 24780714]
10. Efficace F, Jacobs M, Pusic A, Greimel E, Piciocchi A, Kieffer JM, EORTC Quality of Life Group Patient Reported Outcome Measurements Over Time In ONcology-PROMOTION Registry. Patient-reported outcomes in randomised controlled trials of gynaecological cancers: investigating methodological quality and impact on clinical decision-making. Eur J Cancer 2014 Jul;50(11):1925-1941. [doi: 10.1016/j.ejca.2014.04.005] [Medline: 24825114]
11. U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. Health Qual Life Outcomes 2006 Oct 11;4:79 [FREE Full text] [doi: 10.1186/1477-7525-4-79] [Medline: 17034633]
12. Snyder CF, Aaronson NK, Choucair AK, Elliott TE, Greenhalgh J, Halyard MY, et al. Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. Qual Life Res 2012 Oct;21(8):1305-1314. [doi: 10.1007/s11136-011-0054-x] [Medline: 22048932]
13. Snyder CF, Herman JM, White SM, Luber BS, Blackford AL, Carducci MA, et al. When using patient-reported outcomes in clinical practice, the measure matters: a randomized controlled trial. J Oncol Pract 2014 Sep;10(5):e299-e306 [FREE Full text] [doi: 10.1200/JOP.2014.001413] [Medline: 24986113]
14. Müller V, Nabieva N, Häberle L, Taran F, Hartkopf AD, Volz B, et al. Impact of disease progression on health-related quality of life in patients with metastatic breast cancer in the PRAEGNANT breast cancer registry. Breast 2018 Feb;37:154-160. [doi: 10.1016/j.breast.2017.08.008] [Medline: 29237546]
15. Triberti S, Savioni L, Sebri V, Pravettoni G. eHealth for improving quality of life in breast cancer patients: A systematic review. Cancer Treat Rev 2019 Mar;74:1-14. [doi: 10.1016/j.ctrv.2019.01.003] [Medline: 30658289]

16. Schwartzberg L. Electronic Patient-Reported Outcomes: The Time Is Ripe for Integration Into Patient Care and Clinical Research. Am Soc Clin Oncol Educ Book 2016;35:e89-e96 [FREE Full text] [doi: 10.1200/EDBK_158749] [Medline: 27249775]
17. Gwaltney CJ, Shields AL, Shiffman S. Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. Value Health 2008;11(2):322-333 [FREE Full text] [doi: 10.1111/j.1524-4733.2007.00231.x] [Medline: 18380645]
18. Mulhern B, O'Gorman H, Rotherham N, Brazier J. Comparing the measurement equivalence of EQ-5D-5L across different modes of administration. Health Qual Life Outcomes 2015 Nov 26;13:191 [FREE Full text] [doi: 10.1186/s12955-015-0382-6] [Medline: 26612336]
19. Matthies LM, Taran F, Keilmann L, Schneeweiss A, Simoes E, Hartkopf AD, et al. An Electronic Patient-Reported Outcome Tool for the FACT-B (Functional Assessment of Cancer Therapy-Breast) Questionnaire for Measuring the Health-Related Quality of Life in Patients With Breast Cancer: Reliability Study. J Med Internet Res 2019 Jan 22;21(1):e10004 [FREE Full text] [doi: 10.2196/10004] [Medline: 30668517]
20. Campbell N, Ali F, Finlay AY, Salek SS. Equivalence of electronic and paper-based patient-reported outcome measures. Qual Life Res 2015 Aug;24(8):1949-1961. [doi: 10.1007/s11136-015-0937-3] [Medline: 25702266]
21. Wallwiener M, Matthies L, Simoes E, Keilmann L, Hartkopf AD, Sokolov AN, et al. Reliability of an e-PRO Tool of EORTC QLQ-C30 for Measurement of Health-Related Quality of Life in Patients With Breast Cancer: Prospective Randomized Trial. J Med Internet Res 2017 Sep 14;19(9):e322 [FREE Full text] [doi: 10.2196/jmir.8210] [Medline: 28912116]
22. Graf J, Moreno B, Wallwiener M, Menzel K, Brucker SY, Simoes E. [Practicability and Efficiency of E-health Applications in Patient-Reported Outcomes: State of and Need for Research]. Gesundheitswesen 2018 Nov;80(11):953-962. [doi: 10.1055/s-0044-100037] [Medline: 29486499]
23. Steele Gray C, Gill A, Khan AI, Hans PK, Kuluski K, Cott C. The Electronic Patient Reported Outcome Tool: Testing Usability and Feasibility of a Mobile App and Portal to Support Care for Patients With Complex Chronic Disease and Disability in Primary Care Settings. JMIR Mhealth Uhealth 2016 Jun 02;4(2):e58 [FREE Full text] [doi: 10.2196/mhealth.5331] [Medline: 27256035]
24. Schäffeler N, Pfeiffer K, Grischke EM, Wallwiener D, Garbe C, Zipfel S, et al. [Acceptance and reliability of an electronic psychooncological screening of patients with breast cancer: a randomized controlled study]. Psychother Psychosom Med Psychol 2013 Sep;63(9-10):374-380. [doi: 10.1055/s-0032-1333301] [Medline: 23420280]
25. Coons SJ, Gwaltney CJ, Hays RD, Lundy JJ, Sloan JA, Revicki DA, ISPOR ePRO Task Force. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. Value Health 2009 Jun;12(4):419-429 [FREE Full text] [doi: 10.1111/j.1524-4733.2008.00470.x] [Medline: 19900250]
26. Hartkopf AD, Graf J, Simoes E, Keilmann L, Sickenberger N, Gass P, et al. Electronic-Based Patient-Reported Outcomes: Willingness, Needs, and Barriers in Adjuvant and Metastatic Breast Cancer Patients. JMIR Cancer 2017 Aug 07;3(2):e11 [FREE Full text] [doi: 10.2196/cancer.6996] [Medline: 28784595]
27. Bennett AV, Jensen RE, Basch E. Electronic patient-reported outcome systems in oncology clinical practice. CA Cancer J Clin 2012;62(5):337-347 [FREE Full text] [doi: 10.3322/caac.21150] [Medline: 22811342]
28. Wintner LM, Giesinger JM, Zabernigg A, Rumpold G, Sztankay M, Oberguggenberger AS, et al. Evaluation of electronic patient-reported outcome assessment with cancer patients in the hospital and at home. BMC Med Inform Decis Mak 2015 Dec 23;15(1):110 [FREE Full text] [doi: 10.1186/s12911-015-0230-y] [Medline: 26699708]
29. Kikawa Y, Hatachi Y, Rumpold G, Tokiwa M, Takebe S, Ogata T, et al. Evaluation of health-related quality of life via the Computer-Based Health Evaluation System (CHES) for Japanese metastatic breast cancer patients: a single-center pilot study. Breast Cancer 2019 Mar;26(2):255-259 [FREE Full text] [doi: 10.1007/s12282-018-0905-1] [Medline: 30194662]

Abbreviations

EORTC QLQ-C30: European Organization for the Research and Treatment of Cancer core quality of life questionnaire
ePRO: electronic patient-reported outcome
ePROCOM: electronic Patient-Reported Outcomes and Compliance Analysis
FACT-B: Functional Assessment of Cancer Therapy-Breast Cancer
HRQoL: health-related quality of life
pPRO: paper-based patient-reported outcome
PRO: patient-reported outcome

Edited by R Kukafka, G Eysenbach; submitted 04.09.19; peer-reviewed by A Birukov, G Pravettoni; comments to author 08.10.20; revised version received 24.01.21; accepted 23.09.21; published 08.02.22.

Please cite as:
Graf J, Sickenberger N, Brusniak K, Matthies LM, Deutsch TM, Simoes E, Plappert C, Keilmann L, Hartkopf A, Walter CB, Hahn M, Engler T, Wallwiener S, Schuetz F, Fasching PA, Schneeweiss A, Brucker SY, Wallwiener M
Implementation of an Electronic Patient-Reported Outcome App for Health-Related Quality of Life in Breast Cancer Patients: Evaluation and Acceptability Analysis in a Two-Center Prospective Trial
J Med Internet Res 2022;24(2):e16128
URL: https://www.jmir.org/2022/2/e16128
doi:10.2196/16128
PMID:35133288

Original Paper

Continuous Glucose Monitoring With Low-Carbohydrate Nutritional Coaching to Improve Type 2 Diabetes Control: Randomized Quality Improvement Program

Dina H Griauzde[1,2,3], MSc, MD; Grace Ling[4], MS; Daniel Wray[5], MBA, MS; Melissa DeJonckheere[3,4], PhD; Kara Mizokami Stout[1,2,6], MS, MD; Laura R Saslow[7], PhD; Jill Fenske[4], MD; David Serlin[4], MD; Spring Stonebraker[4], BS; Tabassum Nisha[4], MS; Colton Barry[4]; Rodica Pop-Busui[2], MD, PhD; Ananda Sen[4,8], PhD; Caroline R Richardson[3,4], MD

[1]VA Ann Arbor Healthcare System, Ann Arbor, MI, United States

[2]Department of Internal Medicine Division of General Medicine, University of Michigan, Ann Arbor, MI, United States

[3]Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, MI, United States

[4]Department of Family Medicine, University of Michigan, Ann Arbor, MI, United States

[5]Twine Clinical Consulting LLC, Park City, UT, United States

[6]Department of Internal Medicine Division of Metabolism, Endocrinology and Diabetes, University of Michigan, Ann Arbor, MI, United States

[7]Department of Health Behavior and Biological Sciences, School of Nursing, University of Michigan, Ann Arbor, MI, United States

[8]Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI, United States

Corresponding Author:
Dina H Griauzde, MSc, MD
VA Ann Arbor Healthcare System
2215 Fuller Rd
Ann Arbor, MI, 48105
United States
Phone: 1 603 860 1066
Email: dhafez@med.umich.edu

Abstract

Background: Type 2 diabetes mellitus (T2DM) is a leading cause of morbidity and mortality globally, with adverse health consequences largely related to hyperglycemia. Despite clinical practice guideline recommendations, effective pharmacotherapy, and interventions to support patients and providers, up to 60% of patients diagnosed with T2DM are estimated to have hemoglobin A_{1c} (HbA_{1c}) levels above the recommended targets owing to multilevel barriers hindering optimal glycemic control.

Objective: The aim of this study is to compare changes in HbA_{1c} levels among patients with suboptimally controlled T2DM who were offered the opportunity to use an intermittently viewed continuous glucose monitor and receive personalized low-carbohydrate nutrition counseling (<100 g/day) versus those who received usual care (UC).

Methods: This was a 12-month, pragmatic, randomized quality improvement program. All adult patients with T2DM who received primary care at a university-affiliated primary care clinic (N=1584) were randomized to either the UC or the enhanced care (EC) group. Within each program arm, we identified individuals with HbA_{1c} >7.5% (58 mmol/mol) who were medically eligible for tighter glycemic control, and we defined these subgroups as UC–high risk (UC-HR) or EC-HR. UC-HR participants (n=197) received routine primary care. EC-HR participants (n=185) were invited to use an intermittently viewed continuous glucose monitor and receive low-carbohydrate nutrition counseling. The primary outcome was mean change in HbA_{1c} levels from baseline to 12 months using an intention-to-treat difference-in-differences analysis comparing EC-HR with UC-HR groups. We conducted follow-up semistructured interviews to understand EC-HR participant experiences with the intervention.

Results: HbA_{1c} decreased by 0.41% (4.5 mmol/mol; *P*=.04) more from baseline to 12 months among participants in the EC-HR group than among those in UC-HR; however, only 61 (32.9%) of 185 EC-HR participants engaged in the program. Among the EC-HR participants who wore continuous glucose monitors (61/185, 32.9%), HbA_{1c} was 1.1% lower at 12 months compared with baseline (*P*<.001). Interviews revealed themes related to EC-HR participants' program engagement and continuous glucose monitor use.

Conclusions: Among patients with suboptimally controlled T2DM, a combined approach that includes continuous glucose monitoring and low-carbohydrate nutrition counseling can improve glycemic control compared with the standard of care.

(J Med Internet Res 2022;24(2):e31184) doi:10.2196/31184

KEYWORDS

type 2 diabetes mellitus; continuous glucose monitoring; low-carbohydrate counseling

Introduction

Background

Type 2 diabetes mellitus (T2DM) is a leading cause of morbidity and mortality globally, with adverse health consequences largely related to hyperglycemia [1]. Unfortunately, despite clinical practice guideline recommendations [2], effective pharmacotherapy [3], and interventions to support patients and providers [4-6], up to 60% of patients with diagnosed T2DM are estimated to have hemoglobin A_{1c} (HbA_{1c}) levels above recommended targets [2,7]. Multilevel barriers hinder optimal glycemic control, including those at the level of patients (eg, medication nonadherence [8]), providers (eg, clinical inertia [9,10]), and health systems (eg, lack of support resources [11]). Such barriers may be exacerbated by the high costs of many T2DM medications, including insulin [12].

Novel strategies that can be sustained and scaled in diverse clinical settings are needed to help more patients with T2DM to achieve the dual goals of glycemic control and reduced medication burden. Growing evidence suggests that these goals are achievable through dietary carbohydrate restriction. Very low–carbohydrate and low-carbohydrate diets (defined as <10% and 10%-26% of total daily energy from carbohydrates, respectively) have been successfully used in clinical trial settings to manage and reverse T2DM [13,14]. Accordingly, clinical practice guidelines for T2DM now support the use of carbohydrate-restricted meal plans among patients with T2DM who (1) are not meeting glycemic targets, (2) wish to reduce the use of antihyperglycemic agent, or (3) prefer such a dietary approach [15,16]. However, to date, few strategies exist to support the use of carbohydrate-restricted meal plans in general practice settings, as such diets often require intensive personalized instruction and close monitoring with medication adjustments to mitigate the risk of hypoglycemia among patients treated with agents other than metformin [17-19].

A promising strategy to effectively, efficiently, and safely support the use of carbohydrate-restricted meal plans among patients with suboptimally controlled T2DM may be through continuous glucose monitoring (CGM). CGM can support patients' self-education and self-management by providing real-time information on individuals' glycemic responses to specific foods. CGM technology—historically used in the management of type 1 diabetes mellitus—is now less expensive, user-friendly, and increasingly used to guide medication treatment decisions among patients with T2DM [20,21]. However, little is known about the potential role of CGM technology as a tool to help patients initiate and sustain behavior changes. In a previously published pilot study of 15 patients with prediabetes, we showed that CGM plus low-carbohydrate coaching is an acceptable and feasible approach to support dietary behavior change [22].

Objectives

We hypothesized that an intervention combining the use of CGM technology with dietitian-delivered education focused on dietary carbohydrate restriction would be effective in altering patients' eating behaviors and improving glycemic control among patients with suboptimally controlled T2DM. The primary objective of this pragmatic randomized quality improvement (QI) program is to compare mean change in HbA_{1c} levels among patients with suboptimally controlled T2DM (defined as HbA_{1c} ≥7.5%) who were offered the opportunity to use a CGM and receive nutritional counseling versus those who received usual care (UC).

Methods

Study Design

We conducted a 12-month, pragmatic, randomized controlled population level QI program evaluation. Although we could have conducted a more traditional, simple 2-arm randomized controlled trial, we wanted to be able to understand the potential reach of the QI program in a typical primary care office setting. Thus, rather than randomizing participants after obtaining consent as done in a usual research study, we first identified and randomized the entire population of adult patients with T2DM seen in the clinic to either a UC or an enhanced care (EC) arm. We then attempted to contact every patient who was randomized to the EC arm who had an HbA_{1c} level >7.5% (58 mmol/mol) and for whom tighter glycemic control was medically appropriate and invited them to engage in the program. This allowed us to estimate the potential reach of the program outside of a consented research setting. All patients who engaged in the study provided clinical consent, but only a subset of those who engaged consented to allow detailed data to be used for subsequent research or education. After 1 year, those who were eligible and randomized to the UC arm were also offered the opportunity to engage in the intervention, but the data presented here do not include the waitlist control results.

The evaluation was pragmatic in that the recruitment and intervention procedures were integrated into usual clinical workflows and were conducted by clinic staff rather than by research staff. We used the Pragmatic Explanatory Continuum Indicator Summary [23] to design the study and select strategies that were more pragmatic (rather than explanatory), thus increasing the likelihood that the program could be scaled and sustained in real-world clinical settings.

Owing to the complex evaluation design, the program was reviewed under 2 separate institutional review board (IRB)

applications by the University of Michigan's IRB. The QI program was deemed exempt and not regulated and was granted a waiver of informed consent for cohort identification and participant tracking (HUM00147295). This component included the quantitative evaluation of the primary outcomes assessed through deidentified cohort summary data. The second IRB application was applied only to the data repository subcohort. All individuals who engaged in the program were invited to participate in a data repository, which allowed study team members to review their complete electronic health record (EHR) for current and subsequent research and education. Participation in the data repository was voluntary and required written informed consent (HUM00148100). The program evaluation was conducted from November 2018 to November 2019.

Participants and Setting

This QI program took place at a university-affiliated family medicine clinic staffed by family medicine physicians, residents, and advanced practice providers. Although detailed patient-level sociodemographic data were limited owing to the QI nature of this program, most patients served by the clinic are White and have private insurance or Medicare.

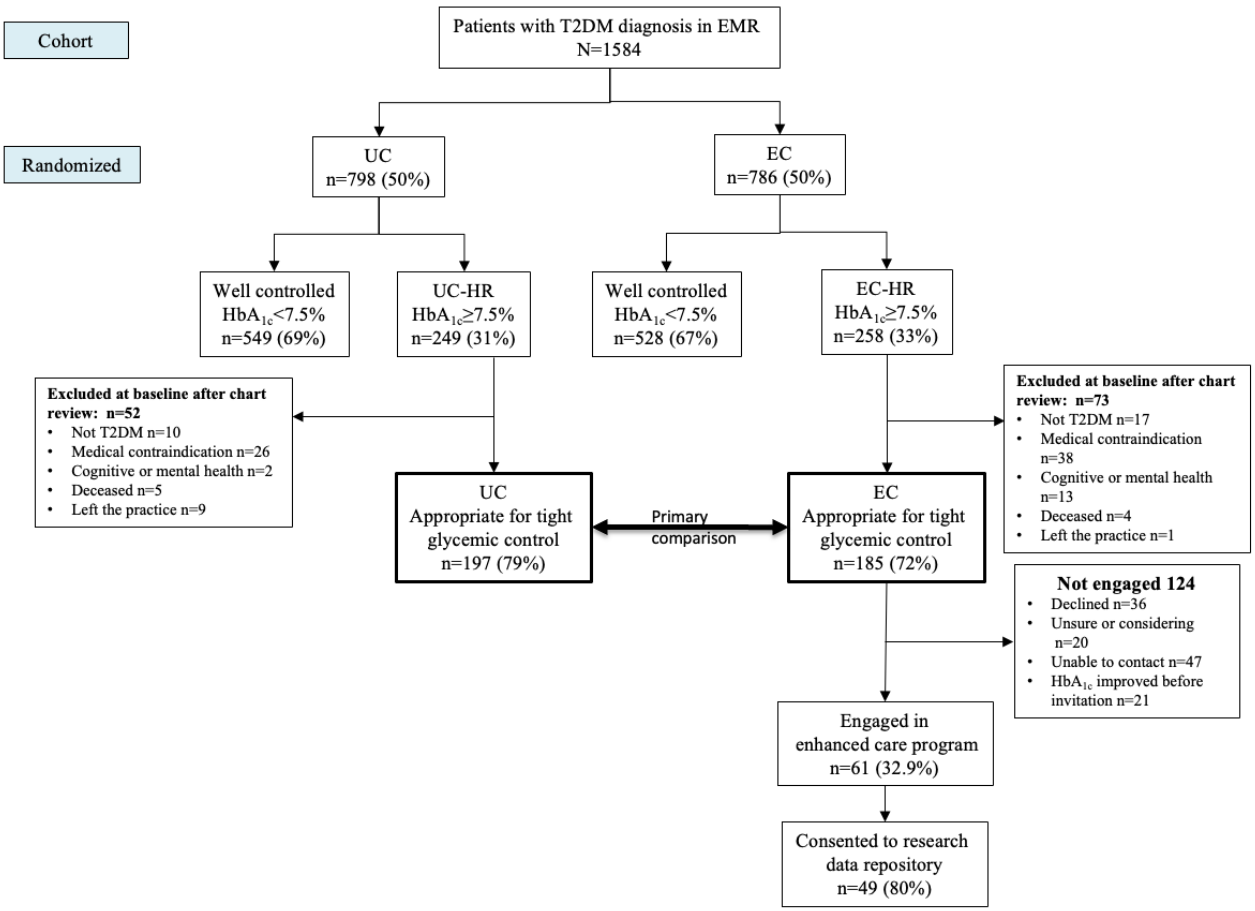
Using Data Direct [24]—the Michigan Medicine web-based tool for accessing data from >4 million patients within the health system—we identified patients aged ≥21 years with T2DM, as determined by EHR-based problem list diagnosis; with HbA_{1c} ≥6.5%; or active prescription for any antihyperglycemic medication other than metformin (N=1584).

Randomization

All individuals were randomized to one of the two program groups using 1:1 randomization with stratification based on age, gender, and BMI: UC or EC. The allocation sequence was generated using STATA 16.0 (StataCorp LLC).

As shown in Figure 1, individuals within each study group were classified as UC-high risk (UC-HR) or EC-HR if they had a baseline HbA_{1c} level ≥7.5% and were medically eligible for improved glycemic control as determined by clinical review of the patients' EHR data and discussion with a primary care physician (PCP), if necessary. Specifically, those individuals for whom tighter glycemic control (ie, HbA_{1c} <7.5%) was not recommended by the American Diabetes Association guidelines [25], older frail individuals at high risk of hypoglycemia and falls, those with a life expectancy of <6 months owing to a comorbid condition, or those with severe or untreated mental health conditions including eating disorders; women who were pregnant or breastfeeding; and those who had previous weight loss surgery were excluded from the high-risk cohorts.

Figure 1. Patients with type 2 diabetes mellitus diagnosis (T2DM) in the electronic medical record (EMR). EC: enhanced care; HbA_{1c}: hemoglobin A_{1c}; HR: high risk; UC: usual care.



Recruitment

Several recruitment strategies were used to engage EC-HR individuals in the program. These included invitation by postal letters, EHR-based electronic messages, and phone calls by a program team member. In addition, for patients who did not respond to this outreach, an attempt was made to engage them face-to-face when they visited the clinic with their PCP or another health care provider.

Program

UC Arm

Individuals within the UC arm received routine care during the 12-month study period. Routine care included PCP follow-up and the opportunity to receive nutrition counseling with a primary care–based dietitian who counseled patients to follow the standard calorie-restricted American Diabetes Association diet without a specific emphasis on dietary carbohydrate restriction [26].

EC Arm

All EC-HR individuals were invited to participate in the program (n=185). Of the 185 individuals, 61 (32.9%) of individuals engaged in the program and 80% (49/61) of individuals provided informed consent to participate in the data repository. Program participants used an intermittently viewed CGM (Abbott Freestyle Libre) and received personalized, low-carbohydrate dietary counseling from the program dietitian. In contrast to other CGM technologies including newer models of Freestyle Libre, the intermittently viewed CGM requires the user to scan the sensor to obtain results and does not have alarms for hypoglycemia or hyperglycemia. We use the term CGM rather than intermittently viewed CGM, as this is the overarching terminology used in consensus guidelines [1].

The sensor was placed by a dietitian and individuals had the opportunity to wear up to 7 CGMs during the 12-month program period. The program participants met one-on-one with the program dietitian. Although the dietitian was hired for program purposes, she did not have any advanced training that would preclude delivery of the intervention by dietitians in other primary care settings. During the first visit, the dietitian conducted a basic nutrition assessment (eg, diet recall) and instructed individuals on how to maintain an accurate food log that was used in conjunction with CGM data to facilitate nutrition coaching and tailored education.

After at least 10 days, the participants returned for a one-on-one visit with the dietitian who reviewed the CGM data and food logs and helped them understand how their dietary carbohydrate intake influenced their blood glucose levels. All the individuals were initially instructed to limit total daily dietary carbohydrate to ≤100 g per day, as this is a clinically relevant, pragmatic, and achievable target for many individuals [13,14]. The participants were subsequently advised to adjust their carbohydrate intake to optimize time in range, defined as blood glucose of 70-180 mg/dL (3.9-10 mmol/L). Specifically, the dietitian met with the patients to review the glucose monitoring data and food logs with the goal of discerning specific foods (eg, bread and bananas) that triggered glucose excursions. She then helped the participants to identify low-carbohydrate alternatives that met their dietary preferences and budget constraints and a lower carbohydrate target was specified (eg, a total of 50 carbohydrates per day) if the patients desired to count carbohydrates. When a lower carbohydrate goal was specified, the participants were instructed on how to count net carbohydrates (ie, total grams of carbohydrates–grams of fiber) to encourage intake of leafy greens and nonstarchy vegetables. In this way, nutrition counseling was tailored to the individual's needs and aimed to accommodate differences in the degree to which individuals may need to restrict dietary carbohydrates to achieve glycemic control.

Although the dietary recommendations were drawn from many publicly available resources, the guiding principles of the program were based on the *Always Hungry* diet developed by Ludwig [27]. The dietitian also educated the patients regarding the potential risks and side effects of carbohydrate restriction, including hypoglycemia; hypotension; and physical symptoms of headache, fatigue, nausea, and constipation. Moreover, the dietitian communicated via the EHR with primary care clinical pharmacists or medical providers to facilitate timely changes in participants' medications to avoid episodes of severe hypoglycemia and hypotension.

Multimedia Appendix 1 shows an example of the handouts used to teach the participants how to count carbohydrates. Multimedia Appendix 2 shows an example of the type of information reviewed during the visit with the dietitian.

Primary Outcome Measure

Baseline and follow-up HbA_{1c} levels were abstracted from the EHR. All the individuals had a baseline HbA_{1c} level obtained as part of routine clinical care before randomization. Follow-up HbA_{1c} data were obtained by the individuals' PCPs as part of routine clinical care. To facilitate complete data availability, the program's dietitian placed orders for HbA_{1c} levels for EC patients on PCPs' behalf; PCPs could approve or cancel, as clinically indicated. The change in HbA_{1c} level was calculated by subtracting participants' baseline HbA_{1c} level from the follow-up HbA_{1c} level.

Secondary Outcome Measure

Baseline and 12-month BMI were abstracted from the EHR. All the individuals had a baseline measurement of weight and height and calculated the BMI as part of routine clinical care before randomization. Follow-up BMI data were obtained by individuals' PCPs as part of routine clinical care; we abstracted the follow-up BMI nearest to the end of the 12-month study period. Change in BMI was calculated by subtracting participants' baseline BMI from follow-up values.

Exploratory and Process Outcomes

Program Engagement

We evaluated the rate of program engagement among the HR-EC cohort. We defined program engagement as the use of at least one CGM and having at least one meeting with the dietitian.

Change in Monthly Cost of Antihyperglycemic Medications

Change in the cost of antihyperglycemic medications was determined by subtracting the total cost of antihyperglycemic medications at baseline from the total cost of antihyperglycemic medications during follow-ups. Among the 61 individuals who engaged in the program, 49 (80%) consented to allow the study team members to review their complete EHR. Of the 49 participants, 1 (2%) participant left the practice, so 48 charts were reviewed for medication cost change data. A study team member reviewed the participants' EHRs and documented the prescribed antihyperglycemic medications and dosages at the start and end of the study period. The cost of individual antihyperglycemic medications was determined using 2017 data from a private insurance claims database, Clinformatics DataMart Database (OptumInsight). Clinformatics DataMart Database contains deidentified claims capturing health care encounters (ie, office visits, outpatient visits, and inpatient visits) for >80 million privately insured adults and children. All the cost data in Clinformatics DataMart Database are standardized to enable comparisons of costs across patients, data sources, and geographic areas in a consistent manner. We also examined the change in monthly cost of antihyperglycemic medications for the subset of patients who were on insulin at baseline.

Change in CGM Metrics

CGM sensor metrics automatically calculated in the CGM report included average blood glucose level, percentage of time above range, percentage of time in range, percentage of time below range, and the number of low glucose events. In the subset of 45 patients in the EC-HR group who engaged in the program, consented data repository participation, and wore at least two sensors during the 12-month program period, we compared CGM sensor metrics using simple bivariate linear regressions with each CGM metric as the dependent variable and a dichotomous time variable to indicate first or last sensor.

Participant Experiences in the EC-HR Group

We conducted semistructured interviews by phone with participants from the EC-HR arm to explore their experiences with participating in the intervention [1]. A semistructured interview guide was selected so that consistent questions were asked across interviews and interviewers, while still allowing each interviewer to ask follow-up questions that were specific to each participant [28]. Phone interviews were selected to enhance participation among individuals who may have difficulty attending in-person. Interviews were conducted by team members trained in semistructured interviewing techniques, with regular feedback from a qualitative methodologist.

Those who agreed to share their data (49/61, 80% of participants in the EC-HR arm) were contacted by phone by the research coordinator and invited to participate. All participants were given up to 2 phone calls for recruitment and voicemails were left when available. Recruitment ended when sufficient interviews were completed to reach thematic saturation. The participants provided verbal consent. The phone interview was audio-recorded and then professionally transcribed.

Statistical Analysis

Measures of central tendency (eg, proportions, means, and SDs) were used for all descriptive analyses. We compared changes in HbA_{1c} levels and BMI using an intention-to-treat difference-in-differences analytic approach using a linear mixed model. The difference-in-differences is the interaction term between a categorical variable denoting the period (eg, before vs after program period) and the study group (eg, UC-HR vs EC-HR). As a sensitivity analysis, we adjusted the models for age and gender. This had little effect on the parameter estimations for the difference-in-differences analysis. Therefore, we present only unadjusted results.

We estimated pre-post changes in HbA_{1c} levels and BMI among the 61 individuals who engaged in the program using a linear mixed model. We estimated pre-post changes in the cost of antihyperglycemic medications among the 79% (48/61) of individuals who consented to data repository participation. We calculated the mean pre-post costs and used paired *t* tests to determine the significance of difference. All analyses were conducted using STATA 16.0.

Qualitative Data Analysis

A total of 3 team members (MD, TN, and CB) trained in qualitative analysis conducted an inductive thematic analysis [29,30]. First, we reviewed and organized all the transcripts using MaxQDA software. MD, TN, and CB independently coded the same transcript and discussed emerging ideas. We created a list of descriptive codes to represent meaningful segments of text; descriptive codes were then applied to the next transcript. We discussed the application of the codes to ensure that codes were being consistently applied by all team members. At this time, additional codes were added, and other codes were revised as needed to clarify the definitions. This process was repeated for the first 3 transcripts. Next, TN and CB independently coded the remaining transcripts. Application of the codes, relationships between the codes, and their meanings were discussed during regular team meetings. After coding, related codes were grouped into themes and summarized using structured summaries, including a narrative description of the theme and all the quotes associated with the theme.

Results

Baseline Characteristics

Demographic characteristics of the complete population (N=1584) were assessed at baseline (Table 1). Slightly less than half of the cohort was female (740/1584, 46.71%), and the mean age was 63.3 (SD 13.1) years. The UC and EC groups were similar at baseline. Among 185 EC-HR participants, 61 (32.9%) were engaged in the program. As shown in Figure 1, among the 124 individuals who did not engage in the study, the most common reasons for nonengagement included inability to contact individuals (47/124, 37.9%) and decline to participate (36/124, 29%).

Table 1. Baseline characteristics of all patients stratified by study group assignment.

Characteristic	All patients	UC[a] group	EC[b] group	UC-HR[c] group	EC-HR[d] group	Engaged in program
Population, n	1584	798	786	197	185	61
Age (years), mean (SD)	63.3 (13.1)	62.9 (12.8)	63.7 (13.4)	60.2 (11)	60 (12.2)	59 (11.9)
Women, n (%)	740 (46.71)	370 (46.4)	370 (47.1)	75 (38.1)	70 (37.8)	28 (46)
Baseline HbA_{1c}[e] level (%), mean (SD)	7.2 (1.5)	7.2 (1.5)	7.2 (1.6)	8.9 (1.4)	9 (1.6)	9 (1.4)
Baseline HbA_{1c} level (mmol/mol), mean (SD)	55.2 (16.4)	55.2 (16.4)	55.2 (17.5)	73.8 (15.3)	74.9 (17.5)	74.9 (15.3)
Baseline BMI (kg/m^2), mean (SD)	34.6 (7.1)	34.8 (7.1)	34.3 (7.1)	35.5 (6.7)	35.8 (7.2)	37.3 (8.5)

[a]UC: usual care.

[b]EC: enhanced care.

[c]UC-HR: UC-high risk.

[d]EC-HR: EC-high risk.

[e]HbA_{1c}: hemoglobin A_{1c}.

Primary Outcome

Baseline data were collected from all the patients. Of the 61 patients, 50 (82%) patients underwent at least one additional HbA_{1c} evaluation as part of routine clinical care during the 12-month study period. The mean time to follow-up HbA_{1c} level was 262 (SD 83) days. The HbA_{1c} level decreased by 0.41% (4.5 mmol/mol; *P*=.04) more from baseline to 12 months among EC-HR participants than among UC-HR participants. Adjusting for age and gender had little impact on the difference. In the pre–post comparison among the EC-HR participants who wore CGMs (n=61), HbA_{1c} level was 1.1% (12 mmol/mol) lower at 12 months compared with baseline (*P*<.001). The pre–post change in BMI in these participants was not statistically significant (within-group difference −0.6, *P*=.06; Table 2).

Table 2. Pre-post analysis[a] for HbA_{1c}[b] and BMI of program participants at 12 months compared with baseline (N=61).

Characteristic	Baseline, mean (SEM[c])	12-month, mean (SEM)	Difference within group at 12 months	*P* value (2-tailed test)
HbA_{1c} (%)	9 (0.11)	7.9 (0.12)	−1.1	<.01
HbA_{1c} (mmol/mol)	74.9 (1.2)	62.8 (1.3)	−12.1	<.01
BMI	37.3 (0.5)	36.7 (0.51)	−0.6	.06

[a]Values predicted from the mixed model not adjusting for age or gender.

[b]HbA_{1c}: hemoglobin A_{1c}.

[c]SEM: SE of the mean.

Secondary Outcome

Baseline BMI was calculated for all the patients included in the cohort. Of the 61 patients, 53 (87%) patients had at least one additional BMI calculated as part of routine clinical care during the 12-month study period. The mean change in time to follow-up BMI was 287 (SD 81) days. There was no significant difference in BMI change from baseline to 12 months between the EC-HR and UC-HR participants (Table 3). Adjusting for age and gender had little impact on the difference. There was a trend toward modest weight loss among the EC-HR participants who wore CGMs (n=61) from baseline to 12 months, but this change was not statistically significant (−0.6 kg/m^2; *P*=.06).

Table 3. Difference-in-differences analysis[a] for HbA_{1c}[b] and BMI at 12 months compared with baseline [1].

Characteristic and group	Baseline, mean (SEM[c])	12-month, mean (SEM)	Difference within group at 12 months	*P* value (2-tailed test)	Difference-in-differences	*P* value
HbA_{1c} (%)					−0.41	.04
UC-HR[d]	8.9 (0.11)	8.43 (0.12)	−0.47	.001		
EC-HR[e]	9.01 (0.11)	8.12 (0.12)	−0.88	<.001		
HbA_{1c} (mmol/mol)					−4.5	.04
UC-HR	73.8 (1.2)	68.6 (1.3)	−5.2	.001		
EC-HR	75 (1.2)	65.2 (0.12)	−9.6	.001		
BMI					−0.11	.63
UC-HR	35.45 (0.49)	34.89 (0.49)	−0.56	.001		
EC-HR	35.84 (0.5)	35.17 (0.51)	−0.67	<.001		

[a]Values predicted from mixed model not adjusting for age or gender.

[b]HbA_{1c}: hemoglobin A_{1c}.

[c]SEM: SE of the mean.

[d]UC-HR: UC-high risk.

[e]EC-HR: EC-high risk.

Exploratory and Process Outcomes

Change in Monthly Cost of Antihyperglycemic Medication

The average change in cost of antihyperglycemic medications from baseline to 12 months was –US $107 (SE of the mean 129.7; *P*=.41) among the 79% (48/61) of individuals who wore sensors and consented to data repository participation. Among the 52% (25/48) of individuals who used insulin, the average change in monthly cost of antihyperglycemic medications from baseline to 12 months was −$363 (SE of the mean 227.1; *P*=.12).

Change in CGM Sensor Metrics

Participants in the EC-HR group who engaged in the program and consented to data repository participation wore an average of 4.3 (SD 1.8) sensors during the 12-month program period. Of the 48 individuals, 4 (8%) individuals wore only 1 sensor. Of the 94% (45/48) of participants who wore at least two sensors, the average glucose decreased (−29.1 mg/dL, SD 9.4 mg/dL; *P*=.003), the average percentage of time above range decreased (−19%, SD 5.8%; *P*=.002), and the average percentage of time in range increased (17.7%, SD 5.4%; *P*=.002). The average percentage of time below range (+0.3%, SD 1%; *P*=.86) and the average number of low glucose events (+0.2 events, SD 1 event; *P*=.84) did not change significantly.

Adverse Events

There were no major adverse events. Among the 61 participants, the most commonly reported events included skin irritation (6/61, 10%) or pain at the sensor site (3/61, 5%). Most endorsed transient symptoms did not preclude subsequent CGM use. Patients on oral anticoagulants noted bruising at the sensor site. Of the 61 participants, 1 (2%) patient reported an episode of sensor-detected hypoglycemia, which was determined to be owing to sensor error when she presented to the emergency department.

Qualitative Results

Overview

Of the 61 participants in the EC-HR arm, 21 (34%) participated in semistructured interviews. Thematic analysis resulted in three themes related to EC-HR participants' program engagement and CGM use: (1) ability to understand how specific foods impact blood glucose trends, (2) ease of following a low-carbohydrate diet, and (3) ease of blood glucose monitoring.

Ability to Understand How Specific Foods Impact Blood Glucose Trends

Participants in the EC-HR arm expressed that by using the CGM for blood glucose monitoring and reviewing their CGM data with the dietitian, they learned and better understood how different foods affected their blood glucose levels. Of the 21 participants, 1 (5%) participant explained:

> *The CGM was really good because it, it helped me just focus on, to see it in real time. I could watch my blood sugar rise, and then, see how high it rise, rose, after I ate somethin' like that. So, it was really cool, I liked it.* [Participant 010]

Participants could easily relate the fluctuations in their glucose levels by looking at the spikes and dips in the graphs generated by the CGM. Many participants observed their glucose levels after eating food to understand how certain food items affect their glucose levels in real time. For example:

> *I really reflected on portion control with that study, because, I noticed—and you know like the types of food I was eating, because, it would jack my sugar my sugar way up or down based on what I was eating. Like a protein shake, even though it was, said it was*

> *low carb and all that, my sugar levels would raise and stay up for quite a while. Versus eatin' like a boiled egg, you know, and a piece of toast. It would be more even. So, I learned like some foods are better for me than other foods are.* [Participant 021]

When reviewing the data independently and in collaboration with the dietitian, the participants could relate their mood, energy levels, concentration, and sleep cycles to the dietary choices they had made. Of the 21 participants, 1 (5%) participant summarized the helpful aspects of the intervention for them:

> *Workin' with [the dietitian], and I think knowing, being able to see the spikes and stuff with what I ate. Oatmeal, I love oatmeal, but it spiked it [glucose level].* [Participant 015]

Finally, owing to their observations, participants described making different dietary choices in anticipation of the impact on their blood glucose. For example, participants reported exchanging foods with higher carbohydrates for those with more protein when their blood glucose levels were high. As 1 (5%) of the 21 participants explained:

> *I mean if it was too high, I would change what I was planning to eat next. So, if I checked it on my way home, and I was gonna pick up dinner, I might change, you know. I was gonna get pizza, but my sugar's high, so instead I'm gonna get chicken.* [Participant 008]

When reviewing their CGM data in real time, the participants were able to respond and prevent subsequent spikes in their glucose levels.

Ease of Following a Low-Carbohydrate Diet

Most EC-HR participants reported that they were able to implement a low-carbohydrate diet as part of the study. They commented that they "do well on it," "no problem," or "it went pretty good." However, some participants reported challenges in implementing and maintaining the diet during the intervention period, including preferences for high-carbohydrate foods, dealing with cravings, and the convenience of highly processed foods. Of the 21 participants, 1 (5%) participant who was able to reduce the number of carbohydrates consumed during the intervention still described the challenges:

> *[The diet was] not a lot of fun. Everything I love happens to be a carbohydrate...* [Participant 013]

Another participant described how difficult it was to find and prepare low-carbohydrate foods when busy or away from home, for example, while traveling for work:

> *I travel for work, so...being able to consistently find something to eat when I'm...getting, when it's easier to go through fast food, than to find something lower carb, higher protein. Right. So that, you know, especially when my schedule's, it's halfway through my day and I end up being someplace I wasn't expecting to be.* [Participant 008]

Though an uncommon experience, 5% (1/21) of the participants noted that feeling "different" or high-maintenance was a barrier to the low-carbohydrate diet during the intervention. They explained:

> *I'm tryin' to hide [the diet]. You know there's, uh, I, I don't tell the whole family. My wife knows, my children know, my boys know, and...I basically don't just pass it around because everybody'll start treatin' you different..."Oh can you eat this though? Can you eat this though?" You know, all that...they'll make somethin' for me that nobody else will eat. So, you know, I'd really rather not be treated like that.* [Participant 002]

As participants continued with the low-carbohydrate diet, many of them reported improvements in mood, concentration, energy, and sleep, and fewer cravings for carbohydrates. Participants described having "more energy when I'm not eatin' carbs or drinkin' caffeinated pop and all that" (Participant 001). However, these improvements were not universal. Some reported difficulty in concentrating and less energy, particularly in the first few days after starting the diet.

When asked about their intentions to maintain the low-carbohydrate diet in the future, more than half of the participants indicated that they would stick to the diet, whereas another 29% (6/21) of participants described making slight modifications. These participants often indicated that they had experienced improvements in their health and well-being while participating in the intervention. For example:

> *I actually am on a low carb diet [still], and I feel more energy 'cause of the low carb diet.* [Participant 003]

Those who made modifications described being less strict with carbohydrate intake. Modifications included consuming more carbohydrates "in moderation" while still trying to "eat healthy" and "cut out junk."

In contrast, 14% (3/21) of the participants reported not continuing to eat a low-carbohydrate diet after experiencing difficulty during the intervention. For example, 5% (1/21) of the participants, who described not reducing carbohydrates intake during the intervention, explained that they found it very restrictive to reduce carbohydrates:

> *That was hard because everything on it is stuff that I eat. I'm not a big fish eater, um...so it was hard, 'cause I know with the carbs, you know, breads and pasta and...that's the kind a stuff I like to eat...[During the intervention] I really didn't change my diet out too much. I just ate less.* [Participant 004]

Ease of Blood Glucose Monitoring

The use of CGM influenced the lifestyle of the participants in many ways. First, participants described how using the CGM eliminated the need for a blood glucose monitor to check their blood glucose levels. For many participants, the CGM was more convenient:

> *It's good for anybody to use anywhere at any time.* [Participant 007]

Furthermore, participants overwhelmingly preferred using the CGM to "poking my fingers all the time." Participants were

frequently irritated with pricking their finger for blood glucose checks because it was not only inconvenient but also painful:

> *The pain with the sticks. And I have short little fat fingers, and it's really hard to grab those sticks and you know so much easier to put the monitor up there and...see my numbers.* [Participant 006]

Similarly, another participant commented that the CGM gives "instant results" instead of having to prick and squeeze the finger to measure the blood glucose.

Second, using the CGM resulted in more frequent blood glucose monitoring among EC-HR participants during the intervention. To illustrate, 5% (1/21) of the participants explained the frequency of their monitoring before the intervention:

> *I occasionally tested my blood sugar...It could go for...one every 3 months, to 2 months. Or was a hit and miss kind of a deal...* [Participant 003]

The participant went on to describe how this behavior changed during the intervention:

> *The changes that I noticed is that I checked my sugar more often. [Before the study] it could have been once every 3 months, once a week. So, there wasn't a scheduled time, just whenever I...thought maybe I should. I would track it. And with the monitor it was right there, handy...It didn't interfere with my life to check the blood sugars.* [Participant 003]

As described earlier, the convenience of the monitor made it easier for participants to incorporate blood glucose monitoring into their lifestyle. Another participant similarly described the laborious process when using the glucometer, rather than the "instant" and "handy" CGM system:

> *[With finger pricking] you got you know, the sensor, or the other thing you gotta, alcohol your finger, poke yourself, get the blood out of there. You know, check it. Alcohol your finger again, you know put everything away, throw away the strip. Uh, you know rather than just puttin' the phone up to your arm, boom it's done.* [Participant 020]

Third, using the CGM provided a more comprehensive picture of their blood glucose levels and blood glucose trends. For example, some participants described the benefit of being able to monitor trends in their blood glucose levels overnight and over time. Of the 21 participants, 1 (5%) participant explained how the additional knowledge about blood glucose trends influenced their diabetes management:

> *We found out also that the, uh...the blood glucose drops during the night, sometimes down in the 40s, which was [when I was] asleep. So, which is, you know, something you wouldn't normally pick up. You can track it more, as to what's been going on. It just gives lots more numbers.* [Participant 017]

Despite these benefits, several participants experienced barriers in using the CGM, including the challenges with the adhesive, skin infection, and the cost of CGM sensors postintervention. Most commonly, the adhesive on the CGM was insufficient, causing the sensor to fall off and then it has to be replaced during a subsequent visit. Although not a common experience, 5% (1/21) of the participants complained of skin infections at the site of the sensor insertion. Finally, some participants expressed that they would like to continue using the CGMs to monitor their blood glucose following the intervention, but were limited by the cost and their insurance:

> *Rather than poking my finger, I totally would go for that [the CGM] in place of the other [glucometer]. But my insurance won't cover it, so I can't continue on.* [Participant 004]

Discussion

Principal Findings

This 12-month pragmatic QI program evaluation examined whether CGM technology—in conjunction with low-carbohydrate nutrition counseling—could reduce HbA_{1c} among patients with suboptimally controlled T2DM. The results show a significantly greater reduction in HbA_{1c} (–0.41%; –4.5 mmol/mol; *P*=.04) among individuals randomized to EC (n=185) compared with individuals randomized to UC (n=197). The improvement in HbA_{1c} level was not accompanied by an increase in the cost of antihyperglycemic medications and was associated with improved CGM metrics. Qualitative results indicated that EC-HR participants positively viewed engagement with the intervention and use of the CGM; however, experiences with the low-carbohydrate diet were more variable.

A recent systematic review and meta-analysis revealed significant improvements in glycemic control at 3 and 6 months among low-carbohydrate diet group participants, but these improvements diminished by 12 months [1]. In contrast, other studies testing the effectiveness of very low-carbohydrate ketogenic diets demonstrate significant and durable improvements in HbA_{1c} levels up to 24 months [2,3]. In this study, although only 32.9% (61/185) of EC participants engaged in the program, we observed a significant reduction in HbA_{1c} at 12 months. Given that this program used low-intensity diet counseling, our results suggest that CGM technology may facilitate adherence to a low-carbohydrate diet.

CGM technology is rapidly advancing and is increasingly recognized as a novel tool to support personalized T2DM management [21]. However, to date, few strategies have examined the role of CGM technology in supporting T2DM management among patients in general practice settings [4]. Among existing interventions targeting individuals with suboptimal glycemic control, many aim to promote medication intensification and adherence to prescribed regimens [4,31]. In contrast to these existing strategies, CGM technology coupled with personalized nutrition counseling may facilitate improved glycemic control through behavior change without medication intensification. A small, nonrandomized study used CGM technology and personalized nutrition counseling to promote a low glycemic index breakfast and demonstrated a reduction in glycemic variability at 2-week follow-up [32]. In this study, we similarly observed improved glycemic control without an increase in antihyperglycemic medication use.

Previous studies have demonstrated the efficacy of low-carbohydrate diets for T2DM management [13,33]. However, few previous studies have demonstrated the effectiveness of low-carbohydrate diets among individuals in real-world settings [34,35]. A primary care–based intervention offered an in-person low-carbohydrate program to a nonrandomized group of patients and demonstrated significant improvements in weight and HbA_{1c} levels over the 13-month study period [34]. A web-based intervention available to the general public similarly demonstrated reductions in HbA_{1c} levels and body weight among a nonrandomized cohort of program completers [35]. We augment these previous findings by demonstrating a significant reduction in HbA_{1c} levels among a randomized cohort of patients using an intention-to-treat analytic approach. Our data suggest that personalized nutrition counseling focused on dietary carbohydrate restriction and guided by individuals' CGM data are more effective than routine PCP and dietitian follow-up among patients with suboptimally controlled T2DM. Intervention participants positively engaged in the intervention, noting the benefits of knowledge and support provided by the dietitian. Although not all participants maintained a low-carbohydrate diet after the intervention, they were able to successfully reduce carbohydrate intake during the intervention and noted improvements in their physical health and well-being.

Among the 185 EC-HR individuals eligible for tight glycemic control, 61 (32.9%) engaged in the program. This is considered a high rate of participation, given that individuals were randomized before the program invitation and all eligible individuals were included in the denominator. Among the 124 individuals who did not engage in the study, the most common reason for nonengagement was the inability to contact eligible individuals (47/124, 37.9%), whereas 29% (36/124) declined to participate in the program and 16.1% (20/124) were uncertain about participation and did not engage during the study period. We did not explore reasons why individuals declined to participate and this is something that could be investigated in future work. To our knowledge, few previous studies have explored barriers to and facilitators of CGM technology as a diabetes self-management tool among patients with T2DM [36,37]. Although CGM technology can empower and motivate patients with type 1 diabetes mellitus who are accustomed to routine self-monitoring of blood glucose to guide insulin dosing [36], CGM technology may enhance diabetes-related distress, which is a known barrier to T2DM self-management [38]. Moreover, carbohydrate-restricted diets may not appeal to some patients' preferred eating patterns, thus underscoring the need for additional personalized nutrition approaches [39,40].

Over the 12-month study period, we observed statistically significant within-group changes in HbA_{1c} levels and BMI among UC-HR and EC-HR individuals. This may reflect contemporaneous changes in clinical practice with a shift away from using obesogenic antihyperglycemic medications (eg, insulin and sulfonylureas) and toward newer agents that may facilitate weight loss and patient compliance (eg, sodium-glucose cotransporter-2 inhibitors and glucagon-like peptide 1 receptor agonists). Consistent with results of previous literature on primary care–based interventions to improve glycemic control among patients with suboptimally controlled T2DM [4], we did not observe a significant between-group change in BMI.

Limitations

First, we recruited individuals from a primary care clinic within a US academic medical center and the program content was delivered by a single dietician; therefore, our results may not be generalizable to other clinics or populations. Second, we did not evaluate outcomes beyond 12 months and were therefore unable to assess long-term changes in glycemic control. Third, as this was a pragmatic QI study that did not require consent for randomization, we used HbA_{1c} levels and BMI data that were obtained for clinical purposes. Therefore, there was variability in the time between the baseline and follow-up assessments. Moreover, we were limited in the type of data we could abstract from the EHR and we did not have the ability to review changes in antihyperglycemic medication use for the complete cohort. Among the EC-HR subset of individuals who provided written informed consent for EHR review, medication intensification did not drive between-group change in HbA_{1c} levels. Follow-up interviews were limited to the EC-HR data repository participants and the experiences of participants in the UC arm were not explored. Finally, we cannot discern the comparative effectiveness of dietary carbohydrate restriction plus CGM versus dietary carbohydrate restriction or CGM alone. However, our results suggest that CGM may facilitate the translation of low-carbohydrate diets into routine clinical practice. This contrasts with previous efficacy and effectiveness studies of dietary carbohydrate restriction, which are often intensive or require study-specific personnel, which may limit their generalizability to routine practice settings.

Conclusions

Many patients with T2DM do not achieve optimal glycemic control [7] despite clinical practice guideline recommendations [15,16] and interventions to support the intensification of patients' antihyperglycemic regimens [4]. There is now growing evidence to support the use of carbohydrate-restricted diets to help patients reduce blood glucose levels and medication use [16,18,41]. However, the degree to which individual patients need to restrict dietary carbohydrates to achieve benefits is unknown [19]. Our findings demonstrate that the use of CGM technology and personalized nutrition counseling focused on dietary carbohydrate restriction can help patients with suboptimally controlled T2DM to improve HbA_{1c} levels without increasing antihyperglycemic medication use. As CGM technology evolves [21] and carbohydrate restriction is increasingly accepted as a powerful tool to support T2DM self-management, this program may be a scalable and sustainable strategy to help and empower patients with T2DM to achieve glycemic control.

Acknowledgments

This project was funded by Twine Clinical Consulting, LLC, which is a consultancy organization focused on fostering quality improvement in health care. It has no financial relationship or conflicts of interest with Abbott Laboratories or any other manufacturer of continuous glucose monitoring systems. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors would like to thank Erin Kaleba and Michelle Bochinski at the University of Michigan Data Office for developing the algorithm to identify and follow the quality improvement program cohort and Lark Spire at the University of Michigan institutional review board-Med for helping them navigate the complex institutional review board issues related to program evaluation and data repository.

The data sets used and analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

DG wrote the original draft of the manuscript. GL curated the data. DW contributed to the funding acquisition and conceptualization. MD conducted the methodology and wrote the original draft. KM contributed to the conceptualization and investigation process. JF provided resources and supervision. LS contributed to the conceptualization. DS also contributed to the conceptualization and resources. SS provided project administration and data acquisition. TN, AS, and CB performed the formal analyses. RP provided supervision and conceptualization. CR contributed to the conceptualization and formal analysis.

Conflicts of Interest

CR is editor-in-chief for JMIR Diabetes.

Multimedia Appendix 1
Method to calculate net carbohydrates.
[DOCX File , 262 KB - jmir_v24i2e31184_app1.docx]

Multimedia Appendix 2
An example of the nutrition counseling template used by the program dietitian.
[DOCX File , 15 KB - jmir_v24i2e31184_app2.docx]

References

1. Stratton IM, Adler AI, Neil HA, Matthews DR, Manley SE, Cull CA, et al. Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study. Br Med J 2000 Aug 12;321(7258):405-412 [FREE Full text] [doi: 10.1136/bmj.321.7258.405] [Medline: 10938048]
2. Practice guidelines resources. American Diabetes Association. URL: https://professional.diabetes.org/content-page/practice-guidelines-resources [accessed 2020-01-09]
3. Upadhyay J, Polyzos SA, Perakakis N, Thakkar B, Paschou SA, Katsiki N, et al. Pharmacotherapy of type 2 diabetes: an update. Metabolism 2018 Jan;78:13-42. [doi: 10.1016/j.metabol.2017.08.010] [Medline: 28920861]
4. Murphy ME, Byrne M, Galvin R, Boland F, Fahey T, Smith SM. Improving risk factor management for patients with poorly controlled type 2 diabetes: a systematic review of healthcare interventions in primary care and community settings. BMJ Open 2017 Aug 04;7(8):e015135 [FREE Full text] [doi: 10.1136/bmjopen-2016-015135] [Medline: 28780542]
5. van Eikenhorst L, Taxis K, van Dijk L, de Gier H. Pharmacist-led self-management interventions to improve diabetes outcomes. A systematic literature review and meta-analysis. Front Pharmacol 2017 Dec 14;8:891 [FREE Full text] [doi: 10.3389/fphar.2017.00891] [Medline: 29311916]
6. Chen DY, Wu XV, Chan EY, Goh YS. Nurse-led tele-coaching on modifiable cardiovascular risk factors in people with type 2 diabetes mellitus: a systematic review and meta-analysis. Worldviews Evid Based Nurs 2019 Dec 12;16(6):424-432. [doi: 10.1111/wvn.12409] [Medline: 31721428]
7. Blonde L, Aschner P, Bailey C, Ji L, Leiter LA, Matthaei S, Global Partnership for Effective Diabetes Management. Gaps and barriers in the control of blood glucose in people with type 2 diabetes. Diab Vasc Dis Res 2017 May 01;14(3):172-183 [FREE Full text] [doi: 10.1177/1479164116679775] [Medline: 28467203]
8. Khan H, Lasker SS, Chowdhury TA. Exploring reasons for very poor glycaemic control in patients with Type 2 diabetes. Prim Care Diabetes 2011 Dec;5(4):251-255. [doi: 10.1016/j.pcd.2011.07.001] [Medline: 21782539]
9. Khunti K, Gomes MB, Pocock S, Shestakova MV, Pintat S, Fenici P, et al. Therapeutic inertia in the treatment of hyperglycaemia in patients with type 2 diabetes: a systematic review. Diabetes Obes Metab 2018 Feb 25;20(2):427-437 [FREE Full text] [doi: 10.1111/dom.13088] [Medline: 28834075]
10. Khunti K, Millar-Jones D. Clinical inertia to insulin initiation and intensification in the UK: a focused literature review. Prim Care Diabetes 2017 Feb;11(1):3-12 [FREE Full text] [doi: 10.1016/j.pcd.2016.09.003] [Medline: 27727005]
11. Funnell MM, Bootle S, Stuckey HL. The diabetes attitudes, wishes and needs second study. Clin Diabetes 2015 Jan 16;33(1):32-36 [FREE Full text] [doi: 10.2337/diaclin.33.1.32] [Medline: 25653471]

12. Cefalu WT, Dawes DE, Gavlak G, Goldman D, Herman WH, Van Nuys K, Insulin Access and Affordability Working Group. Insulin access and affordability working group: conclusions and recommendations. Diabetes Care 2018 Jun 08;41(6):1299-1311. [doi: 10.2337/dci18-0019] [Medline: 29739814]
13. Sainsbury E, Kizirian NV, Partridge SR, Gill T, Colagiuri S, Gibson AA. Effect of dietary carbohydrate restriction on glycemic control in adults with diabetes: a systematic review and meta-analysis. Diabetes Res Clin Pract 2018 May;139:239-252. [doi: 10.1016/j.diabres.2018.02.026] [Medline: 29522789]
14. McArdle PD, Greenfield SM, Rilstone SK, Narendran P, Haque MS, Gill PS. Carbohydrate restriction for glycaemic control in Type 2 diabetes: a systematic review and meta-analysis. Diabet Med 2019 Mar 03;36(3):335-348. [doi: 10.1111/dme.13862] [Medline: 30426553]
15. VA/DoD Clinical Practice Guidelines: Management of Diabetes Mellitus in Primary Care (2017). U.S. Department of Veterans Affairs. URL: https://www.healthquality.va.gov/guidelines/CD/diabetes/ [accessed 2019-06-25]
16. American Diabetes Association. 5. Facilitating behavior change and well-being to improve health outcomes: standards of medical care in diabetes-2020. Diabetes Care 2020 Jan 20;43(Suppl 1):48-65. [doi: 10.2337/dc20-S005] [Medline: 31862748]
17. Westman EC, Tondt J, Maguire E, Yancy WS. Implementing a low-carbohydrate, ketogenic diet to manage type 2 diabetes mellitus. Expert Rev Endocrinol Metab 2018 Sep 15;13(5):263-272. [doi: 10.1080/17446651.2018.1523713] [Medline: 30289048]
18. McKenzie AL, Hallberg SJ, Creighton BC, Volk BM, Link TM, Abner MK, et al. A novel intervention including individualized nutritional recommendations reduces hemoglobin A1c level, medication use, and weight in type 2 diabetes. JMIR Diabetes 2017 Mar 07;2(1):e5 [FREE Full text] [doi: 10.2196/diabetes.6981] [Medline: 30291062]
19. Evert AB, Dennison M, Gardner CD, Garvey WT, Lau KH, MacLeod J, et al. Nutrition therapy for adults with diabetes or prediabetes: a consensus report. Diabetes Care 2019 May 18;42(5):731-754 [FREE Full text] [doi: 10.2337/dci19-0014] [Medline: 31000505]
20. Mian Z, Hermayer KL, Jenkins A. Continuous glucose monitoring: review of an innovation in diabetes management. Am J Med Sci 2019 Nov;358(5):332-339. [doi: 10.1016/j.amjms.2019.07.003] [Medline: 31402042]
21. Olczuk D, Priefer R. A history of continuous glucose monitors (CGMs) in self-monitoring of diabetes mellitus. Diabetes Metab Syndr 2018 Apr;12(2):181-187. [doi: 10.1016/j.dsx.2017.09.005] [Medline: 28967612]
22. Yost O, DeJonckheere M, Stonebraker S, Ling G, Buis L, Pop-Busui R, et al. Continuous glucose monitoring with low-carbohydrate diet coaching in adults with prediabetes: mixed methods pilot study. JMIR Diabetes 2020 Dec 16;5(4):e21551 [FREE Full text] [doi: 10.2196/21551] [Medline: 33325831]
23. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. Br Med J 2015 May 08;350:h2147. [doi: 10.1136/bmj.h2147] [Medline: 25956159]
24. DataDirect. Office of Research, University of Michigan Medical School. URL: https://research.medicine.umich.edu/our-units/data-office-clinical-translational-research/self-serve-data-tools/datadirect [accessed 2020-07-08]
25. American Diabetes Association. 6. Glycemic targets: standards of medical care in diabetes-2018. Diabetes Care 2018 Jan 08;41(Suppl 1):55-64 [FREE Full text] [doi: 10.2337/dc18-S006] [Medline: 29222377]
26. American Diabetes Association. 4. Lifestyle management: standards of medical care in diabetes-2018. Diabetes Care 2018 Jan 08;41(Suppl 1):38-50 [FREE Full text] [doi: 10.2337/dc18-S004] [Medline: 29222375]
27. Ludwig D. Always Hungry?: Conquer Cravings, Retrain Your Fat Cells, and Lose Weight Permanently. New York: Hachette Book Group; 2015.
28. DeJonckheere M, Vaughn LM. Semistructured interviewing in primary care research: a balance of relationship and rigour. Fam Med Community Health 2019 Mar 08;7(2):e000057 [FREE Full text] [doi: 10.1136/fmch-2018-000057] [Medline: 32148704]
29. Braun V, Clarke V. What can "thematic analysis" offer health and wellbeing researchers? Int J Qual Stud Health Well-being 2014 Oct 16;9:26152 [FREE Full text] [doi: 10.3402/qhw.v9.26152] [Medline: 25326092]
30. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
31. Murphy ME, McSharry J, Byrne M, Boland F, Corrigan D, Gillespie P, et al. Supporting care for suboptimally controlled type 2 diabetes mellitus in general practice with a clinical decision support system: a mixed methods pilot cluster randomised trial. BMJ Open 2020 Feb 12;10(2):e032594 [FREE Full text] [doi: 10.1136/bmjopen-2019-032594] [Medline: 32051304]
32. Joshi A, Mitra A, Anjum N, Shrivastava N, Khadanga S, Pakhare A, et al. Patterns of glycemic variability during a diabetes self-management educational program. Med Sci (Basel) 2019 Mar 25;7(3):52 [FREE Full text] [doi: 10.3390/medsci7030052] [Medline: 30934620]
33. Meng Y, Bai H, Wang S, Li Z, Wang Q, Chen L. Efficacy of low carbohydrate diet for type 2 diabetes mellitus management: a systematic review and meta-analysis of randomized controlled trials. Diabetes Res Clin Pract 2017 Sep;131:124-131. [doi: 10.1016/j.diabres.2017.07.006] [Medline: 28750216]
34. Unwin D, Cuthbertson D, Feinman R, Sprung V. A pilot study to explore the role of a low-carbohydrate intervention to improve GGT levels and HbA1c. Diabesity Pract 2015;4(3):102-108 [FREE Full text]

35. Saslow LR, Summers C, Aikens JE, Unwin DJ. Outcomes of a digitally delivered low-carbohydrate type 2 diabetes self-management program: 1-year results of a single-arm longitudinal study. JMIR Diabetes 2018 Aug 03;3(3):e12 [FREE Full text] [doi: 10.2196/diabetes.9333] [Medline: 30291081]
36. Lawton J, Blackburn M, Allen J, Campbell F, Elleri D, Leelarathna L, et al. Patients' and caregivers' experiences of using continuous glucose monitoring to support diabetes self-management: qualitative study. BMC Endocr Disord 2018 Feb 20;18(1):12 [FREE Full text] [doi: 10.1186/s12902-018-0239-1] [Medline: 29458348]
37. Chiu C, Chou Y, Chen Y, Du Y. Impact of new technologies for middle-aged and older patients: in-depth interviews with type 2 diabetes patients using continuous glucose monitoring. JMIR Diabetes 2019 Feb 21;4(1):e10992 [FREE Full text] [doi: 10.2196/10992] [Medline: 30789351]
38. Chew B, Vos R, Metzendorf M, Scholten R, Rutten G. Psychological interventions for diabetes-related distress in adults with type 2 diabetes mellitus. Cochrane Database Syst Rev 2017 Sep 27;9(9):CD011469 [FREE Full text] [doi: 10.1002/14651858.CD011469.pub2] [Medline: 28954185]
39. de Toro-Martín J, Arsenault BJ, Després J, Vohl M. Precision nutrition: a review of personalized nutritional approaches for the prevention and management of metabolic syndrome. Nutrients 2017 Aug 22;9(8) [FREE Full text] [doi: 10.3390/nu9080913] [Medline: 28829397]
40. Gardner CD. Tailoring dietary approaches for weight loss. Int J Obes Suppl 2012 Jul 10;2(Suppl 1):11-15 [FREE Full text] [doi: 10.1038/ijosup.2012.4] [Medline: 25089189]
41. McKenzie A, Hallberg S, Bhanpuri NH, Athinarayanan SJ, Mccue J, McCarter JP, et al. Continuous remote care model utilizing nutritional ketosis improves type 2 diabetes risk factors in patients with prediabetes. Diabetes 2018 Jun 22;67(Supplement 1). [doi: 10.2337/db18-293-OR]

Abbreviations

CGM: continuous glucose monitoring
EC: enhanced care
EC-HR: enhanced care-high risk
EHR: electronic health record
HbA_{1c}: hemoglobin A_{1c}
IRB: institutional review board
PCP: primary care physician
QI: quality improvement
T2DM: type 2 diabetes mellitus
UC: usual care
UC-HR: usual care-high risk

Edited by A Mavragani; submitted 14.06.21; peer-reviewed by P McArdle, M Salimi, G Zimmermann; comments to author 28.08.21; revised version received 16.09.21; accepted 30.11.21; published 02.02.22.

Please cite as:
Griauzde DH, Ling G, Wray D, DeJonckheere M, Mizokami Stout K, Saslow LR, Fenske J, Serlin D, Stonebraker S, Nisha T, Barry C, Pop-Busui R, Sen A, Richardson CR
Continuous Glucose Monitoring With Low-Carbohydrate Nutritional Coaching to Improve Type 2 Diabetes Control: Randomized Quality Improvement Program
J Med Internet Res 2022;24(2):e31184
URL: https://www.jmir.org/2022/2/e31184
doi:10.2196/31184
PMID:35107429

Original Paper

Digitally Supported Lifestyle Intervention to Prevent Type 2 Diabetes Through Healthy Habits: Secondary Analysis of Long-Term User Engagement Trajectories in a Randomized Controlled Trial

Piia Lavikainen[1], PhD; Elina Mattila[2], PhD; Pilvikki Absetz[3,4], PhD; Marja Harjumaa[2], PhD; Jaana Lindström[5], PhD; Elina Järvelä-Reijonen[3], PhD; Kirsikka Aittola[3], MSc; Reija Männikkö[3,6], DSc; Tanja Tilles-Tirkkonen[3], DSc; Niina Lintu[7], PhD; Timo Lakka[7,8,9], MD, PhD; Mark van Gils[2], PhD; Jussi Pihlajamäki[3,6], MD, PhD; Janne Martikainen[1], PhD

[1]School of Pharmacy, University of Eastern Finland, Kuopio, Finland

[2]VTT Technical Research Centre of Finland Ltd, Espoo, Finland

[3]School of Medicine, Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio, Finland

[4]Faculty of Social Sciences, Tampere University, Tampere, Finland

[5]Department of Public Health and Welfare, Finnish Institute for Health and Welfare, Helsinki, Finland

[6]Endocrinology and Clinical Nutrition, Department of Medicine, Kuopio University Hospital, Kuopio, Finland

[7]Institute of Biomedicine, University of Eastern Finland, Kuopio, Finland

[8]Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio, Finland

[9]Foundation for Research in Health Exercise and Nutrition, Kuopio Research Institute of Exercise Medicine, Kuopio, Finland

Corresponding Author:
Piia Lavikainen, PhD
School of Pharmacy
University of Eastern Finland
P.O. Box 1627
Kuopio, 70211
Finland
Phone: 358 407024682
Email: piia.lavikainen@uef.fi

Abstract

Background: Digital health interventions may offer a scalable way to prevent type 2 diabetes (T2D) with minimal burden on health care systems by providing early support for healthy behaviors among adults at increased risk for T2D. However, ensuring continued engagement with digital solutions is a challenge impacting the expected effectiveness.

Objective: We aimed to investigate the longitudinal usage patterns of a digital healthy habit formation intervention, BitHabit, and the associations with changes in T2D risk factors.

Methods: This is a secondary analysis of the StopDia (Stop Diabetes) study, an unblinded parallel 1-year randomized controlled trial evaluating the effectiveness of the BitHabit app alone or together with face-to-face group coaching in comparison with routine care in Finland in 2017-2019 among community-dwelling adults (aged 18 to 74 years) at an increased risk of T2D. We used longitudinal data on usage from 1926 participants randomized to the digital intervention arms. Latent class growth models were applied to identify user engagement trajectories with the app during the study. Predictors for trajectory membership were examined with multinomial logistic regression models. Analysis of covariance was used to investigate the association between trajectories and 12-month changes in T2D risk factors.

Results: More than half (1022/1926, 53.1%) of the participants continued to use the app throughout the 12-month intervention. The following 4 user engagement trajectories were identified: terminated usage (904/1926, 46.9%), weekly usage (731/1926, 38.0%), twice weekly usage (208/1926, 10.8%), and daily usage (83/1926, 4.3%). Active app use during the first month, higher net promoter score after the first 1 to 2 months of use, older age, and better quality of diet at baseline increased the odds of

belonging to the continued usage trajectories. Compared with other trajectories, daily usage was associated with a higher increase in diet quality and a more pronounced decrease in BMI and waist circumference at 12 months.

Conclusions: Distinct long-term usage trajectories of the BitHabit app were identified, and individual predictors for belonging to different trajectory groups were found. These findings highlight the need for being able to identify individuals likely to disengage from interventions early on, and could be used to inform the development of future adaptive interventions.

Trial Registration: ClinicalTrials.gov NCT03156478; https://clinicaltrials.gov/ct2/show/NCT03156478

International Registered Report Identifier (IRRID): RR2-10.1186/s12889-019-6574-y

(J Med Internet Res 2022;24(2):e31530) doi:10.2196/31530

KEYWORDS

type 2 diabetes; user engagement; digital behavior change intervention; trajectories; habit formation; mobile health

Introduction

Type 2 diabetes (T2D) is currently one of the most prevalent noncommunicable diseases burdening public health globally [1]. It is a progressing disease that compromises health-related quality of life [2,3], introduces severe comorbidities, and increases premature mortality [4,5]. It also has a significant economic impact on individuals, health systems, and societies [6]. However, research shows that T2D could be prevented efficiently by lifestyle improvement [7-9]. This highlights the need for early diabetes prevention. In addition, plausible effects of lifestyle changes are reported to sustain for several years [10,11]. Still, after 20 years of evidence of effectiveness, health care systems struggle to find and implement scalable individual-level lifestyle change support in routine practice.

Digital health interventions may offer a scalable way to prevent T2D by providing early support for healthy behavior improvement among individuals at increased risk of T2D with minimal burden to health systems. A recent review supported findings on the effectiveness of technology-driven T2D prevention interventions in producing short-term (≤6 months) and long-term (≥12 months) weight loss, with the number of behavior change techniques applied positively associated with effectiveness [12]. However, continuous user engagement in digital solutions remains a challenge [13-15], impacting the expected effectiveness. For this, finding out who will actively use and potentially benefit from digital solutions is essential for providing the right care to the right person at the right time, which, in turn, will potentially improve the personalization and optimization of the care.

The BitHabit app was developed in the Finnish Stop Diabetes (StopDia) project to provide adults at increased risk of T2D an easy way to form healthy lifestyle habits [16]. The applicability of the BitHabit app was further evaluated in the Finnish health care system in a 12-month randomized controlled trial (RCT) [16]. The aims of this study were to (1) identify long-term user engagement trajectories of the BitHabit app among the participants, (2) examine predictors of the trajectories, and (3) investigate associations between the trajectories and changes in T2D risk factors during the trial.

Methods

Study Design and Participants

This study is a secondary analysis of the StopDia trial, which was an unblinded parallel RCT (ClinicalTrials, NCT03156478). The participants of the trial were randomly allocated to a digital intervention group (DIGI; n=967), a group combining the digital intervention and face-to-face group coaching (DIGI+GROUP; n=971), or a control group (n=969). Randomization was done using a computerized randomization system with 1:1:1 allocation after baseline examinations. In this study, only data from the digital intervention arms (ie, DIGI and DIGI+GROUP) were used. The RCT study protocol, including the intervention lifestyle objectives and outcome measures, has been described in more detail in earlier work [16].

The participants were recruited from several communication channels, such as social media, newspapers, radio, television, websites, health care and social service units, and community pharmacies, by encouraging people to visit the website of the project to fill in the StopDia Digital Screening Tool [17]. The participants were adults aged 18 to 74 years living in 3 provinces of Finland and being at increased risk of T2D as identified with the StopDia Digital Screening Tool including the Finnish Diabetes Risk Score (FINDRISC) developed to estimate the 10-year risk of developing drug-treated T2D [18]. Adults at increased risk of T2D, who scored at least 12 points in the FINDRISC, had repeatedly measured impaired fasting glucose or glucose tolerance, or had a history of gestational diabetes, were invited to participate in the study. Other inclusion criteria were a possibility to use a computer, smartphone, or tablet with internet access; having a mobile phone number; and having adequate Finnish language skills. The exclusion criteria were current type 1 or 2 diabetes, current pregnancy or breastfeeding, having an active cancer, and being less than 6 months from active cancer treatment.

Intervention

Participants in the DIGI and DIGI+GROUP received access to the BitHabit app for the 1-year intervention period in 2017-2019. The BitHabit app was designed in the StopDia project based on principles derived from habit formation theories [19,20] and the Self-Determination Theory [21]. The BitHabit app has been described in more detail earlier [22]. Briefly, the main goal of the app was to help the participants form lifestyle habits that

would help prevent T2D. The app provided the user with a broad selection of behaviors in 13 different categories related to diet, physical activity, sleep, positive mood, stress management, smoking, and alcohol consumption. App use included browsing and selecting context-specific minimum-effort behaviors, reporting performances, and monitoring progress. Automatization of the behaviors into frequently repeated habits was facilitated by the nature of the behaviors (simple, with contextual triggers) as well as reminders for reporting performances.

In addition to using the BitHabit app, DIGI+GROUP members also participated in group coaching consisting of 6 meetings organized in local health care centers, as explained in detail previously [22].

Study Visits

At the first examination visit, after the participants had signed informed consent, clinical measurements, including body weight and height, and waist circumference, were taken by a study nurse in a local health care center. Thereafter, participants received a link to a StopDia Digital Questionnaire with standardized and validated questions in 3 sets to assess diet quality, eating behavior, physical activity, sedentary time, smoking, quality of life, stress, sleep, and several other factors related to T2D risk. The second examination consisted of laboratory measures, including blood glucose, taken in the local health care center. All baseline assessments were repeated at the 12-month follow-up.

After using the app for 1 to 2 months, the DIGI and DIGI+GROUP members were also asked to answer questions on app user experience. They were also asked if they would recommend the BitHabit app to their friends, with possible responses ranging from 0 (not at all likely to recommend) to 10 (very likely to recommend).

Assessments

Baseline Characteristics

At baseline, the participants reported on the following clinical factors: age (categorized as <50, 50-59, and ≥60 years), sex, BMI (kg/m^2), waist circumference (cm), glycated hemoglobin A_{1c} (HbA_{1c}; mmol/mol), FINDRISC, diet quality, and physical activity. Obesity was defined as BMI ≥30 kg/m^2, and abdominal obesity was defined as a waist circumference ≥88 cm for women and ≥102 cm for men [23]. Diet quality was assessed using the Healthy Diet Index that ranges between 0 and 100, and is calculated as the sum of scores for 7 main domains, including meal pattern (0-10 points), grains (0-20 points), fruits and vegetables (0-20 points), fats (0-15 points), fish and meat (0-10 points), dairy (0-10 points), and snacks and treats (0-15 points), with higher scores indicating a healthier diet [24]. Perceived self-efficacy related to nutrition was measured with the Nutrition Emotional Barriers Self-Efficacy Score, with higher scores indicating better emotional barrier self-efficacy [25]. Perceived stress was assessed with the Perceived Stress Scale consisting of 10 questions, with higher scores indicating a higher level of experienced stress [26]. Total physical activity (hours/week) was measured as the sum of leisure time total physical activity, combined total physical activity during work trips, and total conditioning and functional physical activity.

Socioeconomic factors considered were education (categorized as elementary school, high or vocational school, and college or academic degree), household size (categorized as single, and two or more members), and household gross income (categorized as <€25,000/year, €25,000-64,999/year, and ≥€65,000/year). A currency exchange rate of €1=US $1.14 is applicable.

Factors describing digital abilities were prior use of healthy lifestyle digital apps and internet use several times per day. The factor describing the intervention was the use of the BitHabit app together with face-to-face group coaching.

Early User Experience and Use Engagement

User experience was measured after using the app for 1 to 2 months as the rating of the likelihood to recommend the app to others (0-10), which was categorized according to the net promoter score definition [27] into detractors (0-6), passives (7-8), and promoters (9-10).

Furthermore, short-term user engagement was measured as the number of app usage days during the first month of use.

User Engagement

Automatically collected log files of BitHabit app use included time-stamped user interactions from which usage metrics, such as usage days, usage sessions and their durations, selected habits, and habit performances, were derived.

Outcomes

In this study, 12-month changes in selected T2D risk factors (ie, diet quality measured with the Healthy Diet Index, waist circumference, BMI, and HbA_{1c}) were used as outcome measures, which were measured as differences (in absolute scale or in percentage) between baseline and the end of the intervention at 12 months.

Statistical Analyses

Latent class growth models can be utilized to identify homogeneous subpopulations when studying developmental trajectories from a heterogeneous population [28-30]. This data-driven modeling technique was applied to cluster participants into distinct trajectories of app user engagement. App user engagement was measured as monthly usage days during 2 to 12 months after initiation of app use. This provided 11 repeated measurements on the number of app usage days within 1 month (0-31 days/month). First month usage days were not included in the trajectory analyses but were used as a predictor for trajectory membership in later analyses, as early use is a known predictor of sustained use [31-33]. The number of classes (ie, trajectories) is *a priori* unknown and must be estimated from the data by iteratively fitting alternative models. Models with 1 to 5 classes and varying shapes for the trajectories (linear, quadratic, and cubic) were fitted. Models were estimated with full-information maximum likelihood, and missing data were not imputed, but all available data were used. The method assumes no variation within the trajectories. Unadjusted analyses were conducted. To find the best model, we utilized information

from fit indices and the classification performance of the model, as well as clinical interpretation of the trajectories. The Bayesian information criterion was used as a measure of model adequacy, with a lower value indicating a better model with optimal balance between complexity and good fit. The Low-Mendel-Rubin likelihood ratio test with a *P* value <.05 was considered to indicate better fit for an *n*-class model than for an *n*−1 class model. Entropy was used to guide in the classification accuracy of the model, with higher values indicating better classification. Small classes with less than 4% of the total population were not accepted, and to ensure at least 70 participants per group, latent class growth modeling was performed using Mplus Version 8 [34].

Differences in baseline characteristics between the participant groups defined by the identified trajectories were examined with the chi-square test for categorical variables and the Kruskal-Wallis test for continuous variables. Thereafter, a multivariable multinomial logistic regression model was applied to examine the factors associated with the trajectory membership. In a sensitivity analysis, the model was also adjusted for belonging to the DIGI+GROUP. User experience was introduced in the main analysis by conducting a subgroup analysis among the participants who responded to the user experience questionnaire after the first 1 to 2 months of use. Finally, analysis of covariance was used to study associations between the trajectories and short-term changes in T2D risk factors. Models were adjusted for age, sex, and the baseline value of the specific T2D risk factor to account for possible differences in baseline values. Analyses were conducted with IBM SPSS Version 25.0 (IBM Corp) and R Version 4.0.4 (R Core Team). Statistical significance was set at *P*<.05.

Ethics Statement

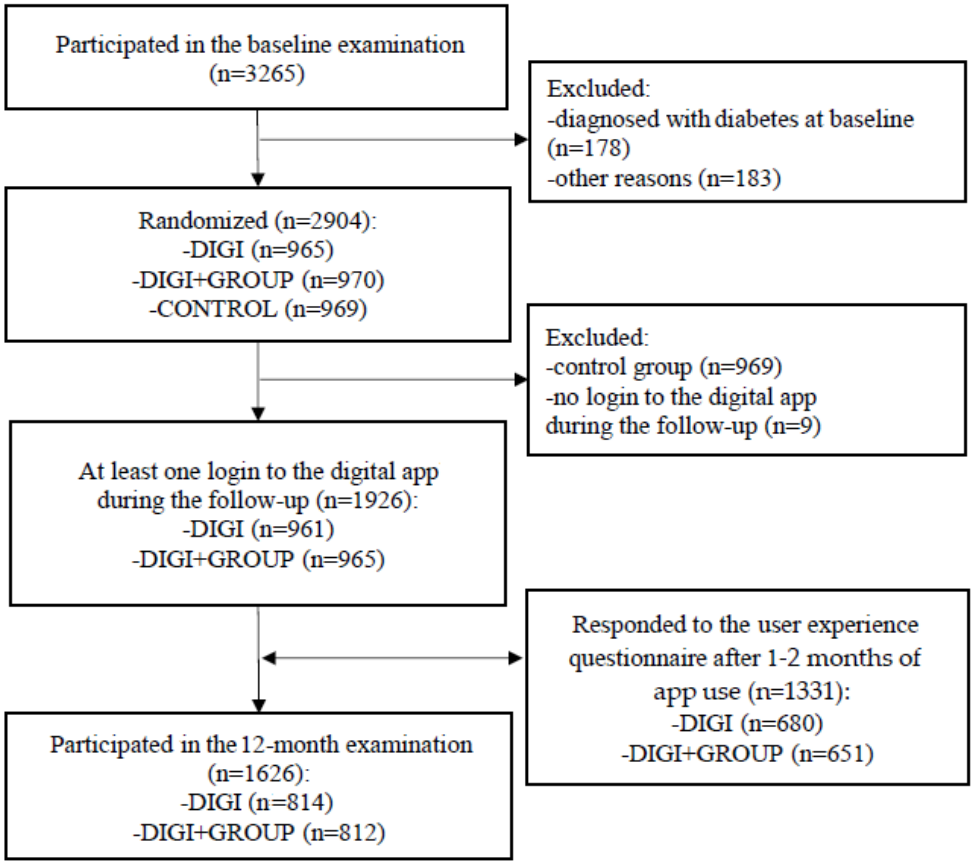
The StopDia study was approved by the Research Ethics Committee of the Hospital District of Northern Savo (number: 467/2016). Written informed consent for participation in the study and the use of data from national health care registers was obtained from all study participants.

Results

Population Characteristics

The study cohort included 1926 participants in the intervention arms with at least one login to the BitHabit app during the study period (Figure 1). Among them, the mean age was 55.2 (SD 10.0) years, and 79.7% (1535/1926) of the participants were women (Multimedia Appendix 1).

Figure 1. Flow chart of the study population. DIGI, digital intervention group; DIGI+GROUP, group combining the digital intervention and face-to-face group coaching.



User Engagement Trajectories

Based on the Low-Mendel-Rubin likelihood ratio test and the clinical interpretation, a 4-class cubic latent class growth model was identified as the best fitting model with good interpretability (Figure 2; Multimedia Appendix 2). Almost half of the participants (904/1926, 46.9%) had very few usage days in the first months of the study, and they dropped close to zero after 6 to 7 months. Another large group (731/1926, 38.0%) started

with approximately 6 usage days per month, which decreased gradually to 3 days per month by the 12th month. A further 10.8% (208/1926) of participants had approximately 14 usage days during the second month with a decrease to 8 days during the 12th month. The remaining 4.3% (83/1926) of participants had almost 23 days of use during the second month, and the usage was sustained with a gradual decrease to 18 days of use during the 12th month.

The identified trajectories were differentiated well when examined with other usage activity metrics (Table 1). Based on the trajectories and all usage metrics, the groups were named as (1) terminated usage, (2) weekly usage, (3) twice weekly usage, and (4) daily usage (Multimedia Appendix 1).

Figure 2. Estimated BitHabit app user engagement trajectories with their 95% CIs.

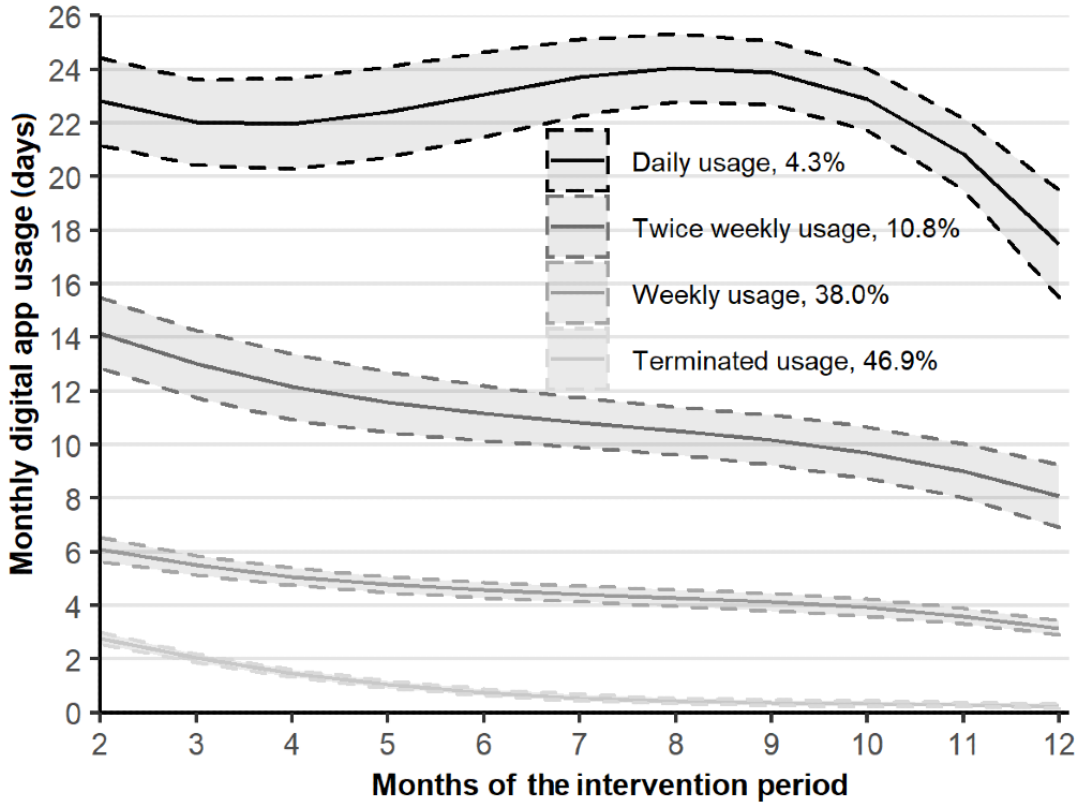


Table 1. Distribution of other use activity metrics by trajectories.

Variable	Terminated usage (n=904), median (IQR)	Weekly usage (n=731), median (IQR)	Twice weekly usage (n=208), median (IQR)	Daily usage (n=83), median (IQR)
Number of usage days	12 (5-23)	54 (44-69)	131 (113-157)	262 (228.25-310)
Number of sessions	14 (6-27)	60 (48-77.75)	153 (128-183.50)	310 (257.25-367.25)
Total duration of use (min)	42.73 (19.55-87.18)	153.63 (89.45-263.72)	302.32 (196.24-509.12)	431.15 (288.45-705.86)
Average session duration (min)	3.23 (1.98-5.23)	2.50 (1.42-4.18)	1.97 (1.21-3.44)	1.32 (0.96-2.02)
Usage period from the first day to the last	124 (51.50-245)	356 (343-360)	360 (354-363)	362 (355.25-364)
Weeks in use	9 (4-17)	40 (35-45)	50 (47.5-51)	51 (50-52)
Months in use	4 (2-7)	12 (12-12)	12 (12-12)	12 (12-12)
Number of selected habits	15 (7-26)	32 (20-53)	43.5 (24-61)	50 (33.5-73)
Number of new habits	7 (2-15)	15 (8-28.75)	26 (13-43)	23 (13-46)
Total number of performances	72 (15-269.5)	720 (352-1470.5)	1433 (804-2562.5)	4516 (2405.5-6281)
Number of habits tracked	13 (6-23.5)	28 (17-46)	39 (22-56.5)	49 (31.5-71.75)

Predictors of User Engagement Trajectories

Active app use during the first month increased the odds of belonging to the more engaged trajectories (odds ratio [OR] 1.65, 95% CI 1.57-1.75 for daily usage; OR 1.45, 95% CI 1.40-1.51 for twice weekly usage; and OR 1.19, 95% CI 1.15-1.22 for weekly usage) than the terminated usage trajectory (Table 2). In addition, older age was associated with increased user engagement compared with the terminated usage category (OR 0.07, 95% CI 0.02-0.19 in those aged <50 years and OR 0.44, 95% CI 0.22-0.87 in those aged 50-59 years for daily usage; OR 0.19, 95% CI 0.11-0.33 in those aged <50 years and OR 0.40, 95% 0.26-0.64 in those aged 50-59 years for twice weekly usage; and OR 0.45, 95% CI 0.34-0.61 in those aged <50 years and OR 0.61, 95% CI 0.47-0.79 in those aged 50-59 years for weekly usage compared with those aged ≥60 years). Women were less likely to be daily users (OR 0.43, 95% CI 0.19-0.98) than terminated users. Better diet quality at baseline, measured with the Healthy Diet Index, increased the odds of belonging to the more engaged trajectories (OR 1.04, 95% CI 1.01-1.08 for daily usage and OR 1.03, 95% CI 1.01-1.05 for twice weekly usage) than the terminated usage trajectory. Prior internet use of several times per day (OR 0.41, 95% CI 0.21-0.79 for daily usage and OR 0.73, 95% CI 0.57-0.95 for weekly usage) decreased the odds of belonging to the more engaged trajectories than the terminated usage trajectory. According to the sensitivity analysis, face-to-face group coaching decreased the odds of belonging to the more engaged trajectories (OR 0.28, 95% CI 0.15-0.52 for daily usage; OR 0.49, 95% CI 0.34-0.73 for twice weekly usage; and OR 0.64, 95% CI 0.52-0.80 for weekly usage) than the terminated usage trajectory (Multimedia Appendix 3). However, there were no more differences between the sexes in trajectory membership.

In a subgroup analysis among 1314 participants who responded to the questionnaire after the first 1 to 2 months of use and had complete data on predictors, app usage days during the first month remained a predictor for trajectory membership (OR 1.56, 95% CI 1.47-1.66 for daily usage; OR 1.37, 95% CI 1.31-1.43 for twice weekly usage; and OR 1.13, 95% CI 1.10-1.17 for weekly usage compared with terminated usage; Multimedia Appendix 4). Furthermore, the detractors classified by the net promoter score had significantly lower odds of belonging to the more engaged trajectories than the terminated usage trajectory (OR 0.23, 95% CI 0.10-0.55 for daily usage; OR 0.19, 95% CI 0.11-0.35 for twice weekly usage; and OR 0.37, 95% CI 0.25-0.55 for weekly usage) compared with promoters of the app.

Table 2. Results of multinomial logistic regression analysis.

Variable	Weekly usage[a] (n=715), aOR[b] (95% CI)	Twice weekly usage[a] (n=204), aOR (95% CI)	Daily usage[a] (n=82), aOR (95% CI)
Age			
<50 years	0.45 (0.34-0.61)[c]	0.19 (0.11-0.33)[c]	0.07 (0.02-0.19)[c]
50-59 years	0.61 (0.47-0.79)[c]	0.40 (0.26-0.64)[c]	0.44 (0.22-0.87)[c]
≥60 years (reference)	1	1	1
Women	0.94 (0.72-1.24)	0.73 (0.44-1.21)	0.43 (0.19-0.98)[c]
Obesity	0.87 (0.70-1.08)	0.63 (0.43-0.92)[c]	0.59 (0.32-1.07)
Healthy Diet Index	1.01 (1.00-1.02)[c]	1.03 (1.01-1.05)[c]	1.04 (1.01-1.08)[c]
Education			
Elementary school	0.77 (0.50-1.18)	0.96 (0.48-1.95)	0.68 (0.22-2.07)
High or vocational school	1.02 (0.79-1.31)	1.32 (0.85-2.06)	1.33 (0.67-2.63)
College or academic degree (reference)	1	1	1
Household size			
Single	0.92 (0.69-1.23)	0.98 (0.57-1.68)	1.53 (0.69-3.42)
≥2 members (reference)	1	1	1
Household income per year			
≤€24,999	1.45 (0.97-2.16)	1.03 (0.48-2.22)	2.79 (0.86-9.03)
€25,000-64,999	1.57 (1.21-2.03)[c]	1.66 (1.05-2.63)[c]	2.87 (1.31-6.32)[c]
≥€65,000 (reference)	1	1	1
Prior use of health lifestyle digital apps	1.04 (0.83-1.30)	1.18 (0.79-1.75)	0.95 (0.50-1.80)
Internet use several times per day	0.73 (0.57-0.95)[c]	0.69 (0.44-1.08)	0.41 (0.21-0.79)[c]
App usage days during the first month	1.19 (1.15-1.22)[c]	1.45 (1.40-1.51)[c]	1.65 (1.57-1.75)[c]

[a]Terminated usage (n=895) as a reference.

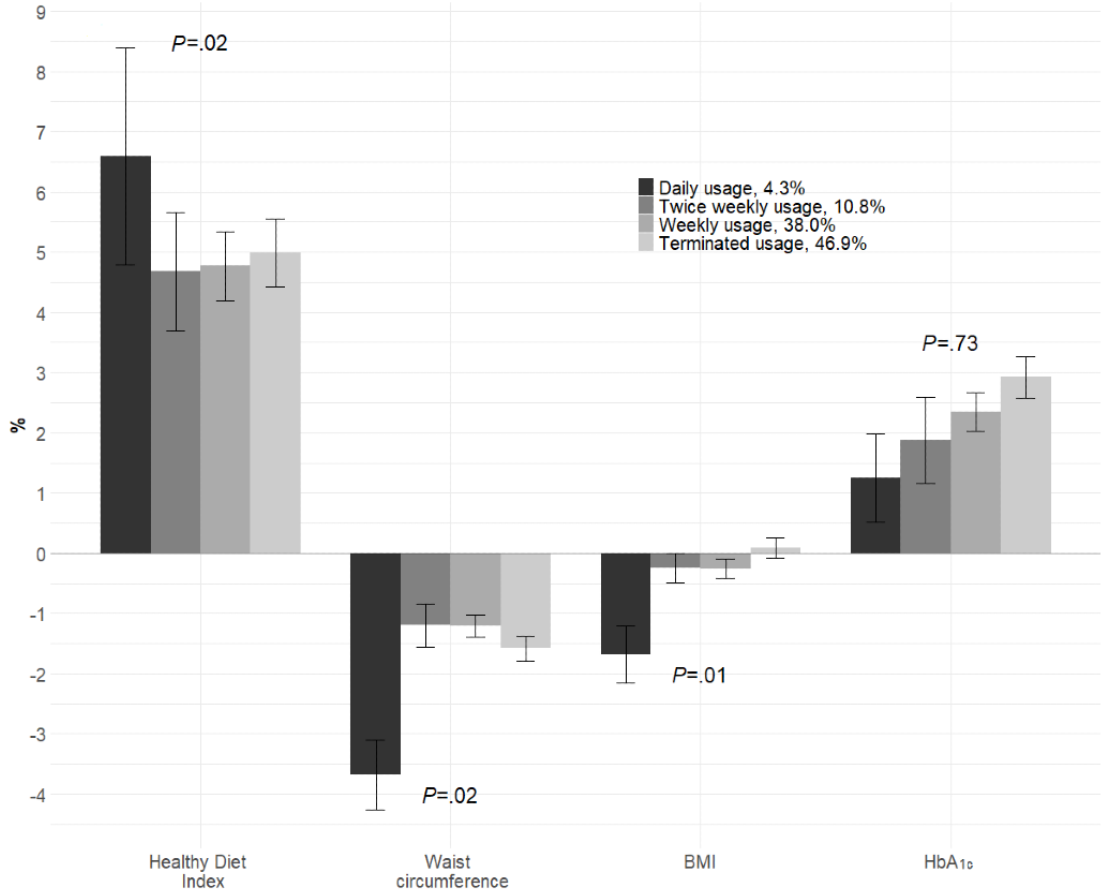
[b]aOR: adjusted odds ratio.

[c]$P<.05$.

Trajectories and Changes in T2D Risk Factors Over 12 Months

High user engagement was associated with 12-month changes in T2D risk factor levels (Figure 3; Multimedia Appendix 5). Compared with the other trajectories, daily usage was associated with a higher increase in the Healthy Diet Index ($P=.02$) and a more pronounced decrease in BMI ($P=.01$) and waist circumference ($P=.02$) at 12 months.

Figure 3. Changes (%) in risk factor levels over 12 months by user engagement trajectories. The error bars represent 95% CIs for means. *P* values are obtained from the analysis of covariance for the main effect of the user engagement trajectory on the change score adjusted for age, sex, and baseline value. HbA_{1c}: glycated hemoglobin A_{1c}.



Discussion

Principal Findings

Four different trajectories of digital app use were identified with a data-driven technique. We observed that initial user engagement during the first month of the intervention, the net promoter score after the first 1 to 2 months of use, older age, and better diet quality at baseline predicted long-term user engagement. While those who had sustained use once to a couple of times per week showed small beneficial changes in risk factor levels in comparison to those who terminated use, those who used the app almost on a daily basis throughout the study showed the most beneficial changes.

Comparison With Prior Research

According to our previously reported results, 99% of participants who were allocated to use the BitHabit app logged into the app at least once and selected at least one habit, and 95% reported at least one habit performance during a 6-month follow-up [22]. No significant differences were observed in cumulative usage metrics between the DIGI and DIGI+GROUP, and more than 50% of the participants in both intervention groups continued to use the app at least on a weekly basis in the first 6 months [22]. This study showed that these active users continued to use the app for the entire 12-month intervention. However, according to the sensitivity analysis, receiving face-to-face group coaching in addition to the digital intervention slightly decreased engagement with the app. Participating in the face-to-face coaching may have decreased the need for later digital support. It should be noted that membership in the DIGI+GROUP arm may have mediated the effects in the sensitivity analysis, and therefore, it was omitted from the main analysis.

Overall, BitHabit app use remained remarkably high in comparison with the use of other digital health apps [22]. For example, Goh et al identified 3 short-term 8-week trajectories of weekly use activity for a caloric-monitoring mobile app among T2D patients [35]. A large proportion of the “minimal user” participants (79%) stopped use within the first 2 weeks. Although women seemed to be more active users, we found no gender differences in app use. Older participants were more likely to show frequent use. These background characteristics have been previously reported to be associated with more active use by others as well [36-38]. This result is promising as it demonstrates that digital interventions can be provided to people of all ages. One can only speculate on the reasons for this finding, but a plausible explanation is higher relevance of T2D prevention. Another explanation could be having less work- and family-related daily activities and hence more time for app use. Other than age and household income, no background factors were found to be associated with app use trajectories. However, prior internet use several times per day predicted

lower user engagement with the app. Intensive internet use has been associated with lower physical activity in adults [39], which may partly explain the observed association in this study.

App use during the first month and the net promoter score were predictors for user engagement trajectories. Previous studies have shown that short-term treatment adherence predicts future adherence outcomes better than baseline characteristics alone, for example, in the case of adherence to self-management for chronic diseases [31,32]. In a recent study, usage during the first 24 hours after login predicted long-term engagement with a weight loss app [33]. We tested the early adherence hypothesis with a 4-week period, considering it a sufficiently long window of opportunity for monitoring and supporting user engagement in real-world prevention programs.

Theoretically, terminating app use does not necessarily imply disengagement from using the behavior change techniques of the intervention [40]. For example, the habit formation approach in the BitHabit app could have led to termination of app use if participants felt that they had already formed sufficiently strong habits or had learnt to follow the habit formation process without the app. Based on the literature, habit formation takes 10 weeks on average, so even for forming multiple habits, 12 months is a rather long usage time [41]. However, our empirical findings suggest that this was not the case in our study, as those who terminated app use had selected less habits than the other user groups and showed no benefits of the intervention in terms of risk factor levels. Moreover, the benefits increased significantly when the frequency of use increased from weekly to daily. Our results are in line with the findings of previous studies reporting that the effectiveness of digital interventions is dependent on user engagement with the intervention [42-44]. Future studies are needed on whether plausible effects of lifestyle interventions sustain the longest among those with the most active engagement during the intervention or whether effective engagement should be determined through other metrics, such as the content of use (eg, type and quality of habits selected and tracked by the users).

The results indicated that only a minor proportion of the participants, belonging to the trajectory of most active users, achieved marked lifestyle changes. Furthermore, the most active users had better well-being at baseline (ie, better diet, more physical activity, and lower stress levels), which is consistent with previous studies [37,38]. Reaching and engaging those with the highest risks and needs are known challenges in most interventions. Adaptive interventions, that is, monitoring for the early indicators of nonresponse and adapting interventions (eg, augmenting or switching interventions), have been proposed as a solution [45]. This seems especially promising in digital interventions, which inherently enable real-time monitoring of user engagement and detection of waning usage as an early indicator of nonresponse. As this study showed, use activity and user experience after the first month of use predicted long-term user engagement. Thus, early use may be regarded as a window of opportunity where user engagement should be monitored and supported especially in real-world prevention programs. If strategies, such as personalized messages and reminders, fail to re-engage participants with the app, they should be offered alternative interventions, such as face-to-face interventions, or other types of digital solutions for weight management and lifestyle changes [46]. Especially in health care that faces cost containment pressures, this kind of approach would help in providing appropriate types of interventions to different individuals.

Strengths and Limitations

The strengths of our study include a real-life setting together with a larger sample size and a longer follow-up than in previous studies on the use of digital health apps in persons with prediabetes or T2D [31,43,47]. Additionally, the studied app could be accessed easily with any device as indicated by an almost 100% first login rate, which reduced selection bias and further strengthened our sample.

However, our study also has some limitations. Women were overrepresented in the sample. The proportion of participants who dropped out during the study and, therefore, the proportion of missing data in outcome variables differed between the identified trajectories. We utilized monthly usage days as a measure for user engagement. However, the sensitivity of this measure in detecting changes in risk factors compared with other user activity metrics remains unknown. The composition of selected habits may also be an important factor for risk factor changes and should be examined in future studies.

Conclusions

Data-driven analysis of digital app user activity is necessary because it reveals true usage patterns. In this study, short-term user engagement, defined as user activity during the first month after the initiation of app use, was found to predict long-term user engagement. Metrics of user experience, based on the net promoter score and measured after early use, was another predictor of long-term user engagement and is a measure that can be easily implemented. These findings could be used in further developments aimed at predicting responses to digital interventions, adapting digital interventions for persons at risk of disengagement, and identifying persons who would benefit the most from using the app.

Acknowledgments

We would like to thank the Strategic Research Council at the Academy of Finland for funding our "Stop Diabetes—from knowledge to solutions (StopDia)" project in 2016-2019 (diary number 303537), the Novo Nordisk Foundation for providing funding in 2018-2020 (diary numbers 33980 and 63753), the Finnish Diabetes Research foundation for providing funding, and the Academy of Finland for funding the project "T2D-DATA" in 2020-2023 (diary numbers 332464 and 332465). We would also like to acknowledge the valuable contributions of the research consortium, our national and international collaborators, the primary health care providers involved in the study, the citizens who participated in the feasibility testing, and the citizens who enrolled

in the study. The funders had no role in designing the study or collecting, managing, or analyzing the data; interpreting the results; writing the manuscript; or deciding to submit the manuscript for publication.

Conflicts of Interest

JM is a founding partner of ESiOR Oy and a board member of Siltana Oy. These companies were not involved in carrying out this research. PL, EM, PA, MH, JL, EJR, KA, RM, NL, TTT, TL, MVG, and JP declare no conflicts of interest.

Multimedia Appendix 1
Baseline characteristics of the study participants in total and by trajectories.
[PDF File (Adobe PDF File), 196 KB - jmir_v24i2e31530_app1.pdf]

Multimedia Appendix 2
Unadjusted latent class growth analyses for app user engagement (monthly usage days during 2-12 months, n=1926).
[PDF File (Adobe PDF File), 55 KB - jmir_v24i2e31530_app2.pdf]

Multimedia Appendix 3
Results of a sensitivity analysis conducted with multivariable multinomial logistic regression.
[PDF File (Adobe PDF File), 159 KB - jmir_v24i2e31530_app3.pdf]

Multimedia Appendix 4
Results of a multivariable multinomial logistic regression analysis among participants in the intervention arms who responded to the user experience questionnaire after the first 1-2 months of use (n=1314).
[PDF File (Adobe PDF File), 171 KB - jmir_v24i2e31530_app4.pdf]

Multimedia Appendix 5
Mean changes in diabetes risk factor levels between baseline and the 12-month follow-up.
[PDF File (Adobe PDF File), 162 KB - jmir_v24i2e31530_app5.pdf]

References

1. IDF Diabetes Atlas. International Diabetes Federation. URL: https://diabetesatlas.org/ [accessed 2022-02-05]
2. Väätäinen S, Keinänen-Kiukaanniemi S, Saramies J, Uusitalo H, Tuomilehto J, Martikainen J. Quality of life along the diabetes continuum: a cross-sectional view of health-related quality of life and general health status in middle-aged and older Finns. Qual Life Res 2014 Sep 8;23(7):1935-1944. [doi: 10.1007/s11136-014-0638-3] [Medline: 24510623]
3. Jalkanen K, Aarnio E, Lavikainen P, Jauhonen H, Enlund H, Martikainen J. Impact of type 2 diabetes treated with non-insulin medication and number of diabetes-coexisting diseases on EQ-5D-5 L index scores in the Finnish population. Health Qual Life Outcomes 2019 Jul 08;17(1):117 [FREE Full text] [doi: 10.1186/s12955-019-1187-9] [Medline: 31286982]
4. Huang Y, Cai X, Mai W, Li M, Hu Y. Association between prediabetes and risk of cardiovascular disease and all cause mortality: systematic review and meta-analysis. BMJ 2016 Nov 23;355:i5953 [FREE Full text] [doi: 10.1136/bmj.i5953] [Medline: 27881363]
5. Nwaneri C, Cooper H, Bowen-Jones D. Mortality in type 2 diabetes mellitus: magnitude of the evidence from a systematic review and meta-analysis. Diabetes & Vascular Disease 2013 Jul 15;13(4):192-207. [doi: 10.1177/1474651413495703]
6. Williams R, Karuranga S, Malanda B, Saeedi P, Basit A, Besançon S, et al. Global and regional estimates and projections of diabetes-related health expenditure: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. Diabetes Res Clin Pract 2020 Apr;162:108072. [doi: 10.1016/j.diabres.2020.108072] [Medline: 32061820]
7. Tuomilehto J, Lindström J, Eriksson JG, Valle TT, Hämäläinen H, Ilanne-Parikka P, Finnish Diabetes Prevention Study Group. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. N Engl J Med 2001 May 03;344(18):1343-1350. [doi: 10.1056/NEJM200105033441801] [Medline: 11333990]
8. Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med 2002 Feb 07;346(6):393-403 [FREE Full text] [doi: 10.1056/NEJMoa012512] [Medline: 11832527]
9. Lindström J, Louheranta A, Mannelin M, Rastas M, Salminen V, Eriksson J, Finnish Diabetes Prevention Study Group. The Finnish Diabetes Prevention Study (DPS): Lifestyle intervention and 3-year results on diet and physical activity. Diabetes Care 2003 Dec 21;26(12):3230-3236. [doi: 10.2337/diacare.26.12.3230] [Medline: 14633807]
10. Lindström J, Peltonen M, Eriksson JG, Ilanne-Parikka P, Aunola S, Keinänen-Kiukaanniemi S, Finnish Diabetes Prevention Study (DPS). Improved lifestyle and decreased diabetes risk over 13 years: long-term follow-up of the randomised Finnish Diabetes Prevention Study (DPS). Diabetologia 2013 Feb 24;56(2):284-293. [doi: 10.1007/s00125-012-2752-5] [Medline: 23093136]

11. Diabetes Prevention Program Research Group. Long-term effects of lifestyle intervention or metformin on diabetes development and microvascular complications over 15-year follow-up: the Diabetes Prevention Program Outcomes Study. The Lancet Diabetes & Endocrinology 2015 Nov;3(11):866-875 [FREE Full text] [doi: 10.1016/S2213-8587(15)00291-0]
12. Van Rhoon L, Byrne M, Morrissey E, Murphy J, McSharry J. A systematic review of the behaviour change techniques and digital features in technology-driven type 2 diabetes prevention interventions. Digit Health 2020 Mar 24;6:2055207620914427 [FREE Full text] [doi: 10.1177/2055207620914427] [Medline: 32269830]
13. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011 Aug 05;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
14. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013 Jan;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
15. Kohl LFM, Crutzen R, de Vries NK. Online prevention aimed at lifestyle behaviors: a systematic review of reviews. J Med Internet Res 2013 Jul 16;15(7):e146 [FREE Full text] [doi: 10.2196/jmir.2665] [Medline: 23859884]
16. Pihlajamäki J, Männikkö R, Tilles-Tirkkonen T, Karhunen L, Kolehmainen M, Schwab U, et al. Digitally supported program for type 2 diabetes risk identification and risk reduction in real-world setting: protocol for the StopDia model and randomized controlled trial. BMC Public Health 2019 Mar 1;19(1):255. [doi: 10.1186/s12889-019-6574-y]
17. Jalkanen K, Järvenpää R, Tilles-Tirkkonen T, Martikainen J, Aarnio E, Männikkö R, StopDia Study Group. Comparison of Communication Channels for Large-Scale Type 2 Diabetes Risk Screening and Intervention Recruitment: Empirical Study. JMIR Diabetes 2021 Sep 09;6(3):e21356 [FREE Full text] [doi: 10.2196/21356] [Medline: 34499036]
18. Lindström J, Tuomilehto J. The diabetes risk score: a practical tool to predict type 2 diabetes risk. Diabetes Care 2003 Mar 01;26(3):725-731. [doi: 10.2337/diacare.26.3.725] [Medline: 12610029]
19. Wood W, Neal DT. Behavioral Science & Policy 2016;2(1):71-83. [doi: 10.1353/bsp.2016.0008]
20. Lally P, van Jaarsveld CHM, Potts HWW, Wardle J. How are habits formed: Modelling habit formation in the real world. Eur. J. Soc. Psychol 2009 Jul 16;40(6):998-1009. [doi: 10.1002/ejsp.674]
21. Ryan R, Deci E. Self-determination theory. Basic psychological needs in motivation, development and wellness. New York, NY: Guilford Press; 2017.
22. Harjumaa M, Absetz P, Ermes M, Mattila E, Männikkö R, Tilles-Tirkkonen T, et al. Internet-Based Lifestyle Intervention to Prevent Type 2 Diabetes Through Healthy Habits: Design and 6-Month Usage Results of Randomized Controlled Trial. JMIR Diabetes 2020 Aug 11;5(3):e15219 [FREE Full text] [doi: 10.2196/15219] [Medline: 32779571]
23. National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation,Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. Circulation 2002 Dec 17;106(25):3143-3421. [Medline: 12485966]
24. Lindström J, Aittola K, Pölönen A, Hemiö K, Ahonen K, Karhunen L, et al. Formation and Validation of the Healthy Diet Index (HDI) for Evaluation of Diet Quality in Healthcare. Int J Environ Res Public Health 2021 Feb 28;18(5):2362 [FREE Full text] [doi: 10.3390/ijerph18052362] [Medline: 33670967]
25. Schwarzer R, Renner B. Social-cognitive predictors of health behavior: Action self-efficacy and coping self-efficacy. Health Psychology 2000;19(5):487-495. [doi: 10.1037/0278-6133.19.5.487]
26. Cohen S. Perceived stress in a probability sample of the United States. In: Spacapan S, Oskamp S, editors. The social psychology of health. Thousand Oaks, CA: Sage Publications; 1988:31-67.
27. Reichheld FF. The one number you need to grow. Harv Bus Rev 2003 Dec;81(12):46-54, 124. [Medline: 14712543]
28. Laird NM, Ware JH. Random-Effects Models for Longitudinal Data. Biometrics 1982 Dec;38(4):963. [doi: 10.2307/2529876]
29. Muthén B, Shedden K. Finite mixture modeling with mixture outcomes using the EM algorithm. Biometrics 1999 Jun;55(2):463-469. [doi: 10.1111/j.0006-341x.1999.00463.x] [Medline: 11318201]
30. Nagin DS, Odgers CL. Group-based trajectory modeling in clinical research. Annu Rev Clin Psychol 2010 Mar;6:109-138. [doi: 10.1146/annurev.clinpsy.121208.131413] [Medline: 20192788]
31. Franklin JM, Krumme AA, Shrank WH, Matlin OS, Brennan TA, Choudhry NK. Predicting adherence trajectory using initial patterns of medication filling. Am J Manag Care 2015 Sep 01;21(9):e537-e544 [FREE Full text] [Medline: 26618441]
32. Kozma CM, Phillips AL, Meletiche DM. Use of an early disease-modifying drug adherence measure to predict future adherence in patients with multiple sclerosis. J Manag Care Spec Pharm 2014 Aug;20(8):800-807. [doi: 10.18553/jmcp.2014.20.8.800] [Medline: 25062073]
33. Andrade AQ, Beleigoli A, Diniz MDF, Ribeiro AL. Influence of Baseline User Characteristics and Early Use Patterns (24-Hour) on Long-Term Adherence and Effectiveness of a Web-Based Weight Loss Randomized Controlled Trial: Latent Profile Analysis. J Med Internet Res 2021 Jun 03;23(6):e26421 [FREE Full text] [doi: 10.2196/26421] [Medline: 34081012]
34. Muthén L, Muthén BO. Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén; 2017.
35. Goh G, Tan NC, Malhotra R, Padmanabhan U, Barbier S, Allen JC, et al. Short-term trajectories of use of a caloric-monitoring mobile phone app among patients with type 2 diabetes mellitus in a primary care setting. J Med Internet Res 2015 Feb 03;17(2):e33 [FREE Full text] [doi: 10.2196/jmir.3938] [Medline: 25648130]

36. Mattila E, Lappalainen R, Välkkynen P, Sairanen E, Lappalainen P, Karhunen L, et al. Usage and Dose Response of a Mobile Acceptance and Commitment Therapy App: Secondary Analysis of the Intervention Arm of a Randomized Controlled Trial. JMIR Mhealth Uhealth 2016 Jul 28;4(3):e90 [FREE Full text] [doi: 10.2196/mhealth.5241] [Medline: 27468653]
37. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun 13;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
38. Edney S, Ryan JC, Olds T, Monroe C, Fraysse F, Vandelanotte C, et al. User Engagement and Attrition in an App-Based Physical Activity Intervention: Secondary Analysis of a Randomized Controlled Trial. J Med Internet Res 2019 Nov 27;21(11):e14645 [FREE Full text] [doi: 10.2196/14645] [Medline: 31774402]
39. Sahin M, Lok S. Relationship between Physical Activity Levels and Internet Addiction of Adults. J Depress Anxiety 2018;07(02):310. [doi: 10.4172/2167-1044.1000310]
40. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and Promoting Effective Engagement With Digital Behavior Change Interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
41. Lally P, Gardner B. Promoting habit formation. Health Psychology Review 2013 May;7(sup1):S137-S158. [doi: 10.1080/17437199.2011.603640]
42. Short C, Rebar A, Plotnikoff R, Vandelanotte C. Designing engaging online behaviour change interventions: a proposed model of user engagement. The European Health Psychologist 2013;17(1):32-38 [FREE Full text]
43. Sepah SC, Jiang L, Ellis RJ, McDermott K, Peters AL. Engagement and outcomes in a digital Diabetes Prevention Program: 3-year update. BMJ Open Diabetes Res Care 2017 Sep 07;5(1):e000422 [FREE Full text] [doi: 10.1136/bmjdrc-2017-000422] [Medline: 28948027]
44. Painter SL, Lu W, Schneider J, James R, Shah B. Drivers of weight loss in a CDC-recognized digital diabetes prevention program. BMJ Open Diabetes Res Care 2020 Jul 13;8(1):e019171 [FREE Full text] [doi: 10.1136/bmjdrc-2019-001132] [Medline: 32624481]
45. Almirall D, Nahum-Shani I, Sherwood NE, Murphy SA. Introduction to SMART designs for the development of adaptive interventions: with application to weight loss research. Transl Behav Med 2014 Sep 6;4(3):260-274 [FREE Full text] [doi: 10.1007/s13142-014-0265-0] [Medline: 25264466]
46. Teeriniemi A, Salonurmi T, Jokelainen T, Vähänikkilä H, Alahäivälä T, Karppinen P, et al. A randomized clinical trial of the effectiveness of a Web-based health behaviour change support system and group lifestyle counselling on body weight loss in overweight and obese subjects: 2-year outcomes. J Intern Med 2018 Nov 04;284(5):534-545 [FREE Full text] [doi: 10.1111/joim.12802] [Medline: 29974563]
47. Alwashmi MF, Mugford G, Abu-Ashour W, Nuccio M. A Digital Diabetes Prevention Program (Transform) for Adults With Prediabetes: Secondary Analysis. JMIR Diabetes 2019 Jul 26;4(3):e13904 [FREE Full text] [doi: 10.2196/13904] [Medline: 31350833]

Abbreviations

DIGI: digital intervention group
DIGI+GROUP: group combining the digital intervention and face-to-face group coaching
FINDRISC: Finnish Diabetes Risk Score
HbA_{1c}: glycated hemoglobin A_{1c}
OR: odds ratio
RCT: randomized controlled trial
T2D: type 2 diabetes

Edited by A Mavragani; submitted 24.06.21; peer-reviewed by A Andrade, Z Huang; comments to author 08.09.21; revised version received 03.11.21; accepted 03.12.21; published 24.02.22.

Please cite as:
Lavikainen P, Mattila E, Absetz P, Harjumaa M, Lindström J, Järvelä-Reijonen E, Aittola K, Männikkö R, Tilles-Tirkkonen T, Lintu N, Lakka T, van Gils M, Pihlajamäki J, Martikainen J
Digitally Supported Lifestyle Intervention to Prevent Type 2 Diabetes Through Healthy Habits: Secondary Analysis of Long-Term User Engagement Trajectories in a Randomized Controlled Trial
J Med Internet Res 2022;24(2):e31530
URL: https://www.jmir.org/2022/2/e31530
doi:10.2196/31530
PMID:35200147

Original Paper

Accessibility of Low-cost Insulin From Illegitimate Internet Pharmacies: Cross-sectional Study

Benjamin Penley[1], PharmD; Lana Minshew[2], PhD; Hui-Han Chen[1], MHS; Stephen Eckel[1], MHA, PharmD; Sachiko Ozawa[1,3], MHS, PhD

[1]Division of Practice Advancement and Clinical Education, Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[2]Center for Innovative Pharmacy Education and Research, Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[3]Department of Maternal and Child Health, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

Corresponding Author:
Sachiko Ozawa, MHS, PhD
Division of Practice Advancement and Clinical Education
Eshelman School of Pharmacy
University of North Carolina at Chapel Hill
CB #7574 Beard Hall 115G
Eshelman School of Pharmacy
Chapel Hill, NC, 27599
United States
Phone: 1 919 966 2626
Email: ozawa@unc.edu

Abstract

Background: There is much public debate regarding the high cost of insulin. With 1-in-4 patients in the United States with type 1 diabetes reporting difficulties affording insulin, there is concern that some of these patients might look for cost savings on the internet, unaware that 96% of internet pharmacies are illegitimate. Patients who purchase insulin from illegitimate internet pharmacies remove themselves from traditional health care systems that ensure safe, quality-assured, and effective medication use.

Objective: This study aims to determine the accessibility of Humalog and NovoLog insulin from internet pharmacies and characterize how these sites approached patient safety, and priced as well as marketed their products.

Methods: From September to December 2019, we queried the phrases *buy insulin online*, *buy Humalog online*, and *buy NovoLog online* in common search engines. The first 100 search results from Google and Bing, and the first 50 search results from Yahoo! and DuckDuckGo were screened. Websites were included if they claimed to sell Humalog or NovoLog insulin, were active, free access, in the English language, and had a unique URL. The legitimacy of websites was classified using LegitScript. Safety and marketing characteristics were compared across the legitimacy of internet pharmacies. Internet pharmacy prices were compared with the prices offered through brick-and-mortar pharmacies using GoodRx.

Results: We found that 59% (n=29) of the 49 internet pharmacies in our analysis were illegitimate, whereas only 14% (n=7) were legitimate and 27% (n=13) were unclassified. Across illegitimate internet pharmacies, Humalog and NovoLog insulin were 2 to 5 times cheaper as compared with both legitimate internet pharmacies and brick-and-mortar stores. Risks associated with the use of illegitimate internet pharmacies by American consumers were evident: 57% (8/14) did not require a prescription, 43% (6/14) did not display medication information or warnings, and only 21% (3/14) offered access to purported pharmacists. This included 9 rogue internet pharmacies that sold Humalog and NovoLog insulin within the United States, where 11% (1/9) required a prescription, 11% (1/9) placed quantity limits per purchase, and none offered pharmacist services. Rogue internet pharmacies often offered bulk discounts (11/18, 61%), assured privacy (14/18, 78%), and promoted other products alongside insulin (13/18, 72%). The marketing language of illegitimate internet pharmacies appealed more to quality, safety, and customer service as compared with legitimate sites.

Conclusions: The ease of access to low-cost insulin through illegitimate internet pharmacies calls for urgent attention. Illegitimate internet pharmacies place patients at risk of poor-quality medications and subpar pharmacy services, resulting in adverse events

and poor diabetes control. A multifaceted approach is needed to close illegitimate internet pharmacies through legal and regulatory measures, develop better search engine filters, raise public awareness of the dangers of illegitimate internet pharmacies, and address the high costs of insulin.

(J Med Internet Res 2022;24(2):e25855) doi:10.2196/25855

KEYWORDS

insulin; diabetes; internet; online; pharmacy; medication; cost

Introduction

Access to Insulin

For patients with type 1 or type 2 advanced diabetes, insulin is the cornerstone of therapy. Furthermore, 1 in 10 Americans (around 34.2 million people) have diabetes, with nearly 1.6 million living with type 1 diabetes [1]. Patients with type 1 diabetes cannot produce endogenous insulin and thus require treatment with exogenous insulin. For patients living with type 2 diabetes, insulin is often required if adequate glycemic control is not maintained with lifestyle modifications and noninsulin medications. In patients who require insulin, regulation of blood glucose is tantamount to disease control; if uncontrolled, it can result in acute, life-threatening conditions of diabetic ketoacidosis or severe hypoglycemia, as well as chronic but still life-threatening complications such as cardiovascular disease, nephropathies, retinopathies, and neuropathies [2]. These complications can result in—among other outcomes—dialysis-dependence, blindness, amputations, serious quality-of-life reductions, and death [2].

Increasingly, rising costs unique to the US market have hindered access to insulin for patients [3]. Despite legislative initiatives to control prescription drug costs, high insulin costs in the United States persist, with list prices of insulin tripling between 2003 and 2013 [3]. Although insulin was first used as a medication in 1922, the insulin market remains dominated by branded products, with no actual generic drug approved. This is important as generic competition has been shown to decrease medicine prices by 60% on average, when 3 or more manufacturers that are generic are in the market [4]. High costs are associated with all insulin types, but rapid-acting insulin analogs such as Humalog (insulin lispro) and NovoLog (insulin aspart) are among the highest priced [3]. Humalog's list price, for example, continued to increase after 2014 from US $391 to nearly US $600 in 2017—a list price increase similar to that of its competitor NovoLog. It is worth noting that net prices remained relatively stable for Humalog and NovoLog during that time frame, meaning that an increase in costs was primarily offset by discounts made available [5]. Regardless, high list prices have a direct impact on patients; in 2019, 1 in 4 patients in the United States with type 1 diabetes reported difficulties affording their medication [6]. Patients such as these are often led to make difficult decisions. More than 1 in 4 patients in the United States with type 1 diabetes reported rationing their insulin in 2019 [7]. Patients who struggle to access insulin from traditional methods, namely brick-and-mortar pharmacies, might look for alternative, lower-cost methods, such as purchasing insulin from friends, across borders, or from illegitimate internet pharmacies [8,9].

Growth of Internet Pharmacies

Internet pharmacies are a popular destination for the purchase of prescription drugs, with 30,000 to 35,000 internet pharmacies accessible in 2016 [10]. Internet pharmacies are defined by whether they operate as legitimate pharmacies or whether they are illegitimate and in violation of US pharmacy laws and practice standards [11]. It has been reported that 96% of all accessible internet pharmacies are illegitimate [10]. According to the World Health Organization, more than 50% of medications acquired from internet pharmacies that do not advertise their physical location (a common characteristic of illegitimate internet pharmacies) are counterfeit or substandard [12]. The dangers associated with illegitimate internet pharmacies have led to the implementation of rules and regulations to ensure safe internet pharmacy use. For example, several states in the United States require internet pharmacies to be accredited with the National Association of Boards of Pharmacy (NABP) to receive licensure [13]. However, the enforcement of such rules and regulations is complicated by an intricate e-commerce environment composed of numerous, often international, stakeholders [11]. The complexity and anonymity of e-commerce allows illegitimate internet pharmacies to avoid detection, and even when detected, reopen operations under new web addresses [14]. Given the evasiveness and high prevalence of illegitimate internet pharmacies, there is concern that patients purchasing medications on the internet might be subject to low-quality products, which could result in the development of dangerous adverse effects, especially for high-risk medications such as insulin.

Beyond concerns related to medication quality, there is also concern regarding the lack of services offered by illegitimate internet pharmacies [15]. In choosing illegitimate internet pharmacies, patients opt out of medication counseling, monitoring, and drug-drug interactions checking that pharmacists and other health care professionals provide to ensure proper medication use [16]. The use of these resources is well documented to improve patient outcomes [17]. Insufficient safety measures could also be further exacerbated by marketing methods that illegitimate internet pharmacies use to attract consumers. Although it has been shown that patients using illegitimate internet pharmacies are at greater risk of developing adverse effects from treatment, there is a lack of current data on how illegitimate internet pharmacies approach patient safety and the marketing methods they use, particularly for high-risk medications [18].

Objective

The accessibility of rapid-acting insulins from illegitimate internet pharmacies could pose a threat to patient safety. We

investigated the availability of Humalog and NovoLog insulin from internet pharmacies through common search engines and documented the website's safety and marketing characteristics, as well as the costs of Humalog and NovoLog insulin.

Methods

Overview

Humalog and NovoLog were chosen for our analysis because of their relatively recent approval by the US Food and Drug Administration (FDA), high list prices, and their outsized role in the discussion of medication pricing in the United States [3]. Website screening was conducted from September to December 2019 using 4 search engines (Google, Bing, Yahoo!, and DuckDuckGo) with the phrases *buy insulin online*, *buy Humalog online*, and *buy NovoLog online*. For each search phrase, the first 100 results from Google and Bing and the first 50 of Yahoo! and DuckDuckGo were screened. Google, Bing, and Yahoo! were chosen for their widespread use in the United States. In addition, DuckDuckGo was searched because of its emphasis on user privacy. Websites were included if they claimed to sell Humalog or NovoLog insulin, were active, free access, in the English language, and had a unique URL. Websites selling Humalog or NovoLog insulin that were accessed from search engine results through a landing page were included. Screenshots were taken of each website for records.

The legitimacy of websites was assessed using LegitScript, which classifies pharmacies according to licensure or registration in affiliated jurisdictions, sale of controlled substances, previous discipline, requirement of valid prescription, protection of privacy, patient services offered, transparency, and domain name registration [19]. LegitScript was chosen to assess internet pharmacy legitimacy because of the breadth of internet pharmacies that it monitors (81,000+) and its partnerships with private (eg, Google, Amazon, Facebook) and governmental agencies (eg, FDA). The websites in this analysis were classified as (1) illegitimate, subclassified as (1a) rogue—*these merchants engage in illegal, unsafe, or misleading activities such as selling prescription drugs without a prescription* and (1b) unapproved—*there is some problem with regulatory compliance or risk, but it is typically less egregious than rogue* or (2) legitimate—*these merchants are registered with a LegitScript certification program and have passed LegitScript certification criteria* or (3) unclassified—no information was available from LegitScript.

The average monthly traffic to website domains defined as unique visits from any country was obtained from SimilarWeb [20]. This website aggregates information on website traffic from a variety of sources. We compared website traffic with the legitimacy of internet pharmacies. The IP addresses of internet pharmacies were examined using IP2location, which retrieves geographic information based on IP addresses [21]. IP address locations were compared with their listed locations.

Costs

Costs of Humalog and NovoLog 100 IU/mL insulin at the most frequently sold dosage forms (ie, 1×10-mL vials, 5×3-mL pens, 5×3-mL cartridges) were collected from a subset of internet pharmacies offering to ship within the United States. Prices per mL of 100 IU/mL Humalog and NovoLog insulin for vials, pens, and cartridges were calculated. The shipping costs and bulk discounts were not considered in the cost calculations. Internet pharmacy prices were compared with prices offered through GoodRx, a drug coupon website. GoodRx prices are representative of out-of-pocket prices that uninsured US patients might pay at brick-and-mortar stores. Average prices at brick-and-mortar stores were obtained from GoodRx on April 30, 2020. Costs were averaged by legitimacy of internet pharmacies among websites selling Humalog and NovoLog pens that offered US shipping.

Marketing

Marketing characteristics were selected based on previous literature, and the analysis focused on internet pharmacies classified by LegitScript [11,16,22,23]. Marketing characteristics were compared across the legitimacy of internet pharmacies. Cost-related marketing characteristics included whether internet pharmacies offered bulk discounts or promo codes. Promotional marketing characteristics included whether internet pharmacies displayed specific medication advertisements pertaining to any form of insulin, advertisements for other products on the page advertising Humalog or NovoLog insulin, or customer testimonies. Marketing characteristics appealing to customer service and general reputability included whether internet pharmacies displayed a phone number, offered the assistance of an associate, or claimed pharmacy registration in some form (eg, Canadian International Pharmacy Association, International Pharmacy Association of British Columbia, or PharmacyChecker [24]). Additional marketing characteristics included whether internet pharmacies offered privacy assurances (eg, discrete packaging or protection of health- or billing-related information) or offered shipping within the United States.

We qualitatively analyzed the website marketing language by collecting texts from the homepages of all included internet pharmacies. After initially reading through the texts to identify the most common marketing language, we selected and defined 6 characteristics. Texts were then screened to assess whether these six characteristics were discussed: (1) quality, (2) safety, (3) customer service, (4) reputability, (5) affordability, and (6) convenience. Quotes representative of each characteristic were collected by legitimacy of internet pharmacies.

Safety

Safety characteristics were selected based on previous literature and focused on internet pharmacies classified by LegitScript [11,16,22,23]. To allow a specific focus on US pharmacies, or those that should be compliant with US laws and regulations, safety characteristics were analyzed only for internet pharmacies that offered shipping within the United States. Safety characteristics were compared across the legitimacy of internet pharmacies. Basic pharmacy-related characteristics included the requirement of a prescription and controls on the amount of Humalog or NovoLog insulin that could be ordered (eg, restricting patients to a 90-day supply or the quantity listed on their prescription). Characteristics related to pharmacy services included whether there was an offer to speak with a pharmacist and whether medication information and drug-related warnings

and precautions were displayed on the product page. Characteristics related to location included whether the pharmacy listed a physical location and the website location listed matched the country of the IP address.

Results

Overview

We screened 300 websites and identified a total of 49 internet pharmacies that claimed to sell Humalog or NovoLog insulin. Of the internet pharmacies, LegitScript classified 59% (29/49) as illegitimate, including 37% (18/49) as rogue and 22% (11/49) as unapproved, whereas 14% (7/49) were legitimate and 27% (13/49) were unclassified. The listed locations of these internet pharmacies differed, with 41% (12/29, all rogue sites) of illegitimate internet pharmacies advertising no location. Of the 29 illegitimate internet pharmacies, 52% (15, 4 rogue and 11 unapproved sites) advertised a Canadian location. The remaining illegitimate internet pharmacies claimed to be located in Great Britain (1/29, 3%) or in Europe (1/29, 3%). No illegitimate internet pharmacies advertised a US location. For legitimate internet pharmacies, 43% (3/7) advertised locations in the United States, 29% (2/7) in Australia, 14% (1/7) in Canada, and 14% (1/7) in India. The majority (4/7, 57%) of legitimate internet pharmacies' physical locations listed on their websites matched those of their server locations. In contrast, physical and server locations matched only 27% (3/11) of the time among unapproved internet pharmacies. None of the physical and server locations matched among the 18 rogue internet pharmacies.

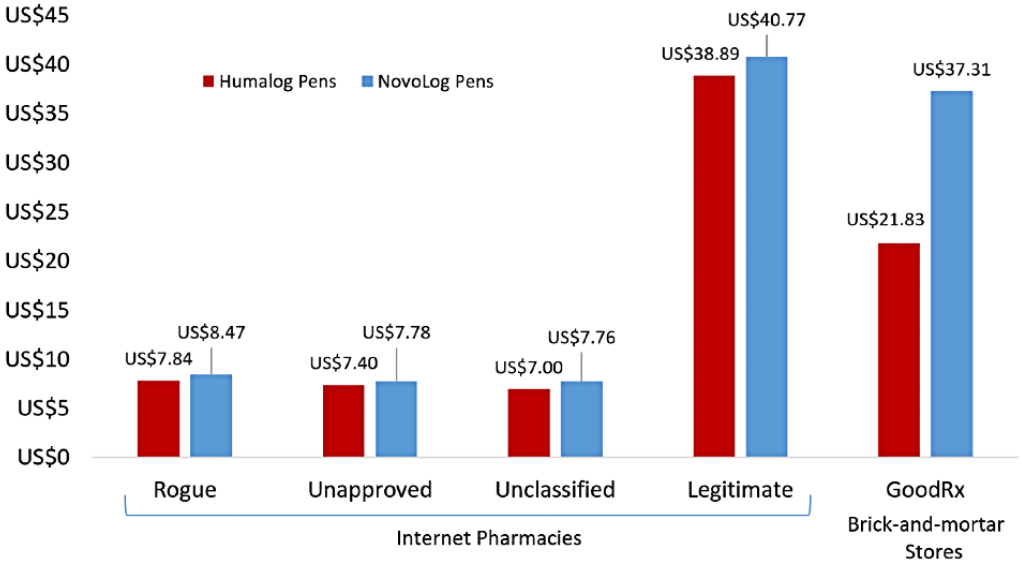
Traffic to internet pharmacies, as determined by SimilarWeb [20], differed depending on the legitimacy of internet pharmacies. Although illegitimate internet pharmacies were the most abundant in the search results, unique monthly visits to each site were comparatively lower for illegitimate internet pharmacies (0-250,000) than to legitimate internet pharmacies (5000-63.6 million). The 3 US-based legitimate internet pharmacies received the highest volumes of unique monthly traffic reported at 63.6 million (Costco), 40.35 million (CVS), and 1.17 million (Healthwarehouse) visits per month [25-27].

Costs

The costs of Humalog and NovoLog insulin vials, pens, and cartridges were recorded only for internet pharmacies shipping within the United States (n=22). The most commonly sold volume and strength of the insulin vials were 10 mL at 100 IU/mL. Humalog and NovoLog insulin 3 mL pens and 3 mL cartridges were most often sold in packages containing a quantity of 5, which is in alignment with product packaging available from brick-and-mortar pharmacies. The cost per mL of 100 IU/mL insulin varied depending on the legitimacy of the internet pharmacies. For rogue internet pharmacies, the average costs of insulin varied depending on dosage forms: Humalog vials cost US $11.30 (n=1), Humalog pens cost US $7.84 (n=4), and Humalog cartridges cost US $10.04 (n=4), whereas NovoLog vials cost US $5.90 (n=2), NovoLog pens cost US $8.47 (n=7), and NovoLog cartridges cost US $8.00 (n=3). Costs were similarly low in unapproved and unclassified internet pharmacies.

For legitimate internet pharmacies, the average costs of insulin (without insurance) were two- to five-fold higher: Humalog vials cost US $31.51 (n=2) and Humalog pens cost US $38.89 (n=2), whereas Humalog cartridges were not available; meanwhile NovoLog vials cost US $32.77 (n=2), NovoLog pens cost US $40.77 (n=2), and NovoLog cartridges cost US $38.47 (n=2). Compared with illegitimate pharmacies, GoodRx costs—representative of average costs at brick-and-mortar pharmacies—were also approximately 2 to 5 times more expensive. However, compared with the cost of legitimate internet pharmacies for uninsured patients excluding shipping costs, GoodRx prices were marginally cheaper: Humalog vials cost US $17.22, Humalog pens cost US $21.83, and Humalog cartridges cost US $34.13, whereas NovoLog vials cost US $29.38, NovoLog pens cost US $37.31, and NovoLog cartridges cost US $35.90. The difference in costs per mL of Humalog and NovoLog insulin pens (the most common dosage form in our analysis) depending on the source is depicted in Figure 1.

Figure 1. Costs of Humalog and NovoLog insulin pens available in the United States on the internet and at brick-and-mortar stores.



Marketing

The marketing characteristics of internet pharmacies selling Humalog and NovoLog insulin and classified by LegitScript (N=36) are described in Table 1. Rogue internet pharmacies differed from unapproved and legitimate internet pharmacies across several characteristics. More often, rogue internet pharmacies offered bulk discounts (11/18, 61%), assured privacy (14/18, 78%), and promoted other products on the Humalog or NovoLog insulin product pages (13/18, 72%). Although rogue internet pharmacies offered some form of contact through email or chat functions, most sites did not offer a phone number (11/18, 61%). Both legitimate (n=7) and unapproved (n=11) internet pharmacies shared similar characteristics, where few offered bulk discounts (2/7, 29% legitimate; 2/11, 18% unapproved), all displayed a phone number (18/18, 100%), and most touted registration or made accreditation claims (6/7, 86% legitimate; 10/11, 91% unapproved).

Table 1. Marketing characteristics of internet pharmacies selling Humalog or NovoLog[a].

Characteristics	Rogue (n=18), n (%)			Unapproved (n=11), n (%)			Legitimate (n=7), n (%)		
	Yes	No	Not reported	Yes	No	Not reported	Yes	No	Not reported
US shipping of insulin	9 (50)	9 (50)	0 (0)	5 (45)	6 (55)	0 (0)	3 (43)	4 (57)	0 (0)
Bulk discounts	11 (61)	5 (28)	2 (11)	2 (18)	8 (73)	1 (9)	2 (29)	5 (71)	0 (0)
Coupons	9 (50)	8 (44)	1 (6)	2 (18)	9 (82)	0 (0)	4 (57)	3 (43)	0 (0)
Registration claims	8 (44)	10 (56)	0 (0)	10 (91)	1 (9)	0 (0)	6 (86)	1 (14)	0 (0)
Privacy assurances	14 (78)	4 (22)	0 (0)	8 (73)	3 (27)	0 (0)	1 (14)	6 (86)	0 (0)
Customer testimonies	10 (56)	8 (44)	0 (0)	7 (64)	4 (36)	0 (0)	3 (43)	4 (57)	0 (0)
Offer to speak with associate	18 (100)	0 (0)	0 (0)	9 (82)	2 (18)	0 (0)	6 (86)	1 (14)	0 (0)
Phone number	7 (39)	11 (61)	0 (0)	11 (100)	0 (0)	0 (0)	7 (100)	0 (0)	0 (0)
Insulin-specific advertisements	0 (0)	18 (100)	0 (0)	3 (27)	8 (73)	0 (0)	0 (0)	7 (100)	0 (0)
Advertisements for other products on page selling insulin	13 (72)	5 (28)	0 (0)	1 (9)	10 (91)	0 (0)	4 (57)	1 (14)	2 (29)

[a]Illustrates marketing characteristics of 36 websites selling Humalog or NovoLog. This table does not include 13 internet pharmacies selling Humalog or NovoLog insulin that were not classified by LegitScript.

Marketing language from website homepages differed according to the legitimacy of internet pharmacies, particularly for quality, safety, and customer service. Although 83% (24/29) of illegitimate internet pharmacies appealed to quality, only 29% (2/7) of legitimate internet pharmacies used language that suggest the quality of medication or services. Marketing language appealing to safety (19/29, 66% illegitimate vs 1/7, 14% legitimate) and customer service (24/29, 83% vs 4/7, 57%) were also more common among illegitimate internet pharmacies. The frequency of use of marketing language was similar among reputability (16/29, 55% vs 4/7, 57%), affordability (24/29, 83% vs 5/7, 71%), and convenience (22/29, 76% vs 5/7, 71%).

Differences in marketing language are also demonstrated through the selection of quotes in Table 2. The marketing language of illegitimate internet pharmacies tended to communicate a sense of urgency to purchase products and strongly emphasized the merits of the pharmacy. For example, one quote from an illegitimate internet pharmacy appealed to reputation, affordability, and safety:

> *If you are looking to buy your prescription drugs in Canada, through a reputable international or online Canadian pharmacy, [our online pharmacy] provides you access to a trusted source of affordable and safe prescription drugs.* [Illegitimate pharmacy]

Table 2. Types of marketing language used on the home pages of internet pharmacies.

Characteristics	Description	Illegitimate (n=29)		Legitimate (n=7)	
		Values, n (%)	Selected quote[a]	Values, n (%)	Selected quote
Quality	Language suggesting quality of medication or services	24 (83)	"We guarantee that all [medications] for sale on this site are 100% genuine and extremely powerful."	2 (29)	"[We] offer you quality care..."
Safety	Language explicitly referring to the safety of medication products, internet ordering platform, or other services	19 (66)	"When it comes to your health, we know that safety is your number one concern. It's ours too."	1 (14)	"Besides delivering medicines at your doorstep, we...help people use their medicines effectively and safely."
Customer service	Language suggesting availability of staff to answer questions or remedy problems	24 (83)	"24/7 customer support (we are always at your disposal!)."	4 (57)	"Customer service: get answers to your questions."
Reputability	Language suggesting renown in selling or accreditation to sell prescription drugs	16 (55)	"[Our pharmacy] has a great reputation serving the community for 47 years and counting."	4 (57)	"Accredited and Certified in all 50 states."
Affordability	Language suggesting discounts or cheap prescription drugs	24 (83)	"[We] provide the same insulin that's available in the US except our prices are much lower, and we pass on the savings to you."	5 (71)	"At [our pharmacy], you can buy health products and medicines online at best discounts."
Convenience	Language suggesting ease of internet pharmacy use or prescription drug purchase; fast delivery or services that allow for time savings	22 (76)	"The process of payment through Bitcoin is simple. You need to go through only a few steps to quickly confirm and complete your order."	5 (71)	"You get the convenience of online shopping combined with the support and guidance of our dedicated team."

[a]Quotes were taken directly from internet pharmacy websites.

Safety

The safety characteristics of internet pharmacies selling Humalog and NovoLog insulin within the United States and classified by LegitScript (N=17) are described in Table 3. Overall, illegitimate internet pharmacies revealed poor patient safety records: 57% (8/14) did not require a prescription, 43% (6/14) did not display medication information or warnings, and only 21% (3/14) offered access to pharmacists. Rogue internet pharmacies differed from legitimate internet pharmacies more substantially than unapproved internet pharmacies in terms of patient safety. Rogue internet pharmacies seldom required a prescription (1/9, 11%) or placed quantity limits on the amount of medication that could be ordered (1/9, 11%), and none offered pharmacist services (0/9, 0%). Unapproved internet pharmacies uniformly claimed to require a prescription (5/5, 100%) and placed quantity limits (5/5, 100%), and some sites offered pharmacist services (3/5, 60%). Drug-related information and warnings were not uniformly displayed for both rogue and unapproved internet pharmacies.

Data were unavailable for one legitimate internet pharmacy regarding whether pharmacist services were offered (where member registration was required). Not all legitimate internet pharmacies were accredited through the NABP because of geography. However, legitimate internet pharmacies required or displayed characteristics consistent with best internet pharmacy communication practices, such as requiring pharmacists to offer individual, meaningful consultations [13].

Table 3. Safety characteristics of internet pharmacies selling Humalog or NovoLog insulin in the United States.

Safety characteristics	Rogue (n=9), n (%)			Unapproved (n=5), n (%)			Legitimate (n=3), n (%)		
	Yes	No	Not reported	Yes	No	Not reported	Yes	No	Not reported
Prescription required	1 (11)	8 (89)	0 (0)	5 (100)	0 (0)	0 (0)	3 (100)	0 (0)	0 (0)
Offer to speak with pharmacist	0 (0)	9 (100)	0 (0)	3 (60)	2 (40)	0 (0)	2 (67)	0 (0)	1 (33)[a]
Medication precautions on product page	6 (67)	3 (33)	0 (0)	2 (40)	3 (60)	0 (0)	3 (100)	0 (0)	0 (0)
Medication information on product page	6 (67)	3 (33)	0 (0)	2 (40)	3 (60)	0 (0)	3 (100)	0 (0)	0 (0)
Quantity control	1 (11)	8 (89)	0 (0)	5 (100)	0 (0)	0 (0)	3 (100)	0 (0)	0 (0)
Lists a physical location	4 (44)	5 (56)	0 (0)	5 (100)	0 (0)	0 (0)	3 (100)	0 (0)	0 (0)
Location listed on website and location of server match	0 (0)	9 (100)	0 (0)	2 (40)	3 (60)	0 (0)	3 (100)	0 (0)	0 (0)

[a]One legitimate pharmacy required an account to gain access to services. Overall, 5 internet pharmacies selling Humalog or NovoLog insulin in the United States were not classified using LegitScript.

Discussion

Principal Findings

Our analysis demonstrates that both Humalog and NovoLog insulin are readily available from internet pharmacies that engage in illegal sales of prescription drugs. Illegitimate internet pharmacies were found to be abundant using common search engines, outnumbering legitimate internet pharmacies. Of the internet pharmacies included in our analysis, nearly 60% (29/49) were illegitimate, whereas only 14% (7/49) were legitimate (with the remainder unclassified). The widespread availability of illegitimate internet pharmacies poses a threat to unsuspecting consumers and provides easy access to those seeking insulin without a prescription [9]. Incentives to purchase insulin from illegitimate internet pharmacies go beyond ease of access, as our analysis reveals that these internet pharmacies offer substantial price reductions as compared with brick-and-mortar stores and legitimate internet pharmacies.

Rising Insulin Costs

In the United States, rising list prices on rapid-acting insulin analogs such as Humalog and NovoLog insulin have resulted in a substantial cost burden for patients with diabetes. Legislative initiatives to curb insulin costs include, among others, setting out-of-pocket maximums and allowing personal drug importation. However, these reforms have only been trialed in some states, and the legality of the personal importation of insulin remains contentious [28,29]. In 2018, in the midst of public outcry at high insulin costs, the manufacturer Eli Lilly introduced Lispro, an authorized generic to Humalog, at a 50% discount on the Humalog list price [30]. Unfortunately, according to a Senate report, the uptake of Lispro has been meager at best, with 83% of 386 surveyed national pharmacies not having Lispro in stock and 69% unable to order the medication [31]. In 2020, the Trump Administration announced that Medicare would begin offering 1750 different insurance plans that capped out-of-pocket spending to US $35 for insulin [32]. However, this out-of-pocket maximum is limited only to Medicare beneficiaries. Given the relatively limited action from the federal government, manufacturers, and other key players in the United States pharmaceutical supply chain, high insulin list prices and out-of-pocket costs for many persist. Many patients continue to face insulin access problems, particularly those who are uninsured, which in 2018 accounted for 8.5% of the US population [3,33].

With financial pressure, some patients who need insulin to manage their diabetes have resorted to illegal activities such as borrowing insulin, importing insulin from lower-cost countries, or purchasing insulin from illegitimate internet pharmacies [8]. Our analysis, which focuses on the internet pharmacy marketplace, demonstrates that among pharmacies that ship within the United States, the per mL cost of Humalog and NovoLog insulin from illegitimate internet pharmacies was approximately 2 to 5 times cheaper than that offered by legitimate internet pharmacies or GoodRx. Such substantial price differences raise concerns that illegitimate internet pharmacies may appeal to patients priced out of traditional means of acquiring insulin. Beyond offering lower prices, rogue internet pharmacies do not require prescriptions and also use marketing methods that appeal to cost-conscious consumers, such as offering bulk discounts or coupons. In addition, illegitimate pharmacies appeal to affordability through language on their homepages.

Although costs are lower for Humalog and NovoLog insulin from illegitimate internet pharmacies, they remain illegal because of the serious risks associated with their use. Humalog and NovoLog insulin are high-risk medications that require both therapeutic monitoring to ensure optimal short- and long-term outcomes and sufficient counseling for the prevention of adverse events such as hypoglycemia. Between 2007 and 2009, nearly 20% of emergency hospitalizations for the treatment of emergent adverse events were because of insulin [34]. The costs of these visits are not trivial; in 2016, the total average cost per person per visit for hypoglycemia was US $1965 for an emergency department visit and US $11,632 for inpatient hospitalization in the United States [35]. The American Diabetes Association recommends that patients treated with insulin who are unaware of hypoglycemia should be counseled on signs and methods to treat it [36].

Threat of Illegitimate Internet Pharmacies

Our analysis demonstrates that illegitimate internet pharmacies, particularly rogue internet pharmacies, do not offer pharmacy services that are on-par with those offered through legitimate internet pharmacies. The majority (8/9, 89%) of rogue internet pharmacies allowed the purchase of Humalog or NovoLog insulin in the United States without a prescription, precluding the involvement of health care professionals in patient care. Further preventing communication with health care professionals, no rogue internet pharmacies made the offer to speak with pharmacists. On illegitimate internet pharmacy websites, patients were often left without medication information and drug-related warnings and precautions. Illegitimate internet pharmacies, with rogue internet pharmacies being the worst offenders, allow patients to access Humalog and NovoLog insulin with minimal information, predisposing these patients to poor diabetes control and potential development of adverse events.

There is also concern as to whether the quality of insulin obtained from illegitimate internet pharmacies is comparable with that obtained from legitimate internet pharmacies. Low costs offered through illegitimate internet pharmacies could suggest less stringent cold chain shipping methods or lower-quality medications. Although our analysis did not collect information on the quality of Humalog and NovoLog insulin obtained from illegitimate internet pharmacies, substandard and falsified medicines are globally prevalent [12,37]. Studies have suggested that poor medication quality from illegitimate pharmacies could further increase patient safety at risk [38]. Illegitimate internet pharmacies, as evinced in our analysis of marketing language, often make claims touting the quality of their products. Ironically, this could amplify the risk of consumers inadvertently purchasing low-quality medications. To address these risks, a multifaceted approach is needed to close illegitimate internet pharmacies, develop better search

engine filters, raise public awareness of the dangers of illegitimate internet pharmacies, and address high insulin costs.

National organizations are combating the proliferation of illegitimate internet pharmacies. The FDA's BeSafeRx campaign and the Alliance for Safe Online Pharmacies' *Buy Safe Rx* campaign help consumers identify illegitimate internet pharmacies and recognize their risks [39,40]. The NABP helps consumers identify legitimate internet pharmacies through their list of accredited digital pharmacies, as well as with a verification service through which legitimate internet pharmacies receive a *dot-pharmacy* domain [13]. The Center for Safe Internet Pharmacies also offers a verification service through its initiative *Verify Before You Buy* (powdered by LegitScript), wherein consumers can search the URL of an internet pharmacy to check its legitimacy [41]. In May 2020, North Carolina's Secretary of State announced a partnership with the Center for Safe Internet Pharmacies to raise consumer awareness [42].

Regulatory and legal actions are ongoing against illegitimate internet pharmacies. Operation Pangea, an effort led by Interpol in conjunction with the FDA and US Department of Justice, has resulted in the removal of thousands of illegitimate internet pharmacies [43]. LegitScript has also collaborated with the FDA to identify and close illegitimate internet pharmacies [44]. However, illegitimate internet pharmacies persist by closing and reopening under new web addresses, requiring continued vigilance by regulatory authorities.

Given that patients with diabetes are frequently counseled on the management of their disease, health care providers such as physicians and pharmacists are uniquely positioned to lead the charge in making consumers aware of the risks of acquiring insulin from illegitimate internet pharmacies.

Limitations

The limitations of our study include the small sample size and the cross-sectional design. However, our screening methods were consistent with what US consumers purchasing insulin on the internet might experience. An additional limitation is that we did not analyze the quality of products available from internet pharmacies, precluding conclusions pertaining to medication quality. We did not purchase medications from these websites because we questioned the ethical implications of financially supporting the operations of illegitimate internet pharmacies. Our analysis was conducted before the COVID-19 pandemic. We believe our results are even more relevant now as internet purchases have become commonplace in the United States [45]. Finally, we were unable to quantify the purchase volume of Humalog or NovoLog insulin from these internet pharmacies. The nonspecific, surrogate measure of unique monthly visits is representative of the overall traffic to these websites and is not necessarily indicative of interest in or purchases of Humalog or NovoLog insulin. Despite these limitations, we contend that our analysis reflects the internet pharmacy marketplace for insulin in the United States as accessed through common search engines.

Conclusions

The relatively low costs of Humalog and NovoLog insulin from easily accessible illegitimate internet pharmacies place patients at risk. Although the elimination of illegitimate internet pharmacies would be the gold standard way to reduce their risk to patients, illegitimate internet pharmacies are elusive. Governmental agencies should continue to pursue legal and regulatory measures with the intent of closing illegitimate internet pharmacies. Search engines should work to filter their results better, decreasing the visibility of illegitimate internet pharmacies. Finally, although public awareness campaigns and provider-to-patient efforts can bring attention to the dangers of illegitimate internet pharmacies, they do not address the reasons that patients may visit these sites. With patient safety in mind, US legislators and members of the pharmaceutical supply chain should work to lower the costs of insulin, thereby diminishing patients' incentive to purchase from illegitimate internet pharmacies.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

The authors thank the directors and coordinators of the Research and Scholarship in Pharmacy (RASP) pathway at the UNC Eshelman School of Pharmacy for their guidance and support. The authors also thank Colleen Higgins on the Global Health Economics for Pharmacy (GHEP) team for her contributions.

Authors' Contributions

BP was involved in the methodology, investigation, data curation, and writing of the original draft; LM was associated with methodology and writing, reviewing, and editing; HHC was involved with methodology, writing, reviewing, and editing; SE was involved with methodology, writing, reviewing, and editing; SO was associated with conceptualization, methodology, supervision, writing, reviewing, and editing.

Conflicts of Interest

None declared.

References

1. Statistics about diabetes. American Diabetes Association. 2018. URL: https://www.diabetes.org/resources/statistics/statistics-about-diabetes [accessed 2021-12-20]

2. Complications. American Diabetes Association. 2021. URL: https://www.diabetes.org/diabetes/complications [accessed 2021-12-20]
3. Cefalu WT, Dawes DE, Gavlak G, Goldman D, Herman WH, Van Nuys K, Insulin AccessAffordability Working Group. Insulin access and affordability working group: conclusions and recommendations. Diabetes Care 2018 Jun;41(6):1299-1311. [doi: 10.2337/dci18-0019] [Medline: 29739814]
4. Dave CV, Hartzema A, Kesselheim AS. Prices of generic drugs associated with numbers of manufacturers. N Engl J Med 2017 Dec 28;377(26):2597-2598. [doi: 10.1056/NEJMc1711899] [Medline: 29281576]
5. Bell J. Lilly reveals Humalog pricing details amid larger scrutiny over insulin costs. BioPharma Dive. 2019. URL: https://www.biopharmadive.com/news/lilly-humalog-price-list-net-insulin/551236/ [accessed 2021-12-20]
6. Herkert D, Vijayakumar P, Luo J, Schwartz JI, Rabin TL, DeFilippo E, et al. Cost-related insulin underuse among patients with diabetes. JAMA Intern Med 2019 Jan 01;179(1):112-114 [FREE Full text] [doi: 10.1001/jamainternmed.2018.5008] [Medline: 30508012]
7. Costs and rationing of insulin and diabetes supplies: findings from the 2018 T1International Patient Survey. T1International. 2018. URL: https://www.t1international.com/media/assets/file/T1International_Report_-_Costs_and_Rationing_of_Insulin__Diabetes_Supplies_2.pdf [accessed 2021-11-22]
8. Litchman ML, Oser TK, Wawrzynski SE, Walker HR, Oser S. The underground exchange of diabetes medications and supplies: donating, trading, and borrowing, oh my!. J Diabetes Sci Technol 2019 Dec 04:1932296819888215 [FREE Full text] [doi: 10.1177/1932296819888215] [Medline: 31801370]
9. Lovett KM, Liang BA, Mackey TK. Online, direct-to-consumer access to insulin: patient safety considerations and reform. J Diabetes Sci Technol 2012 Nov 01;6(6):1503-1506. [doi: 10.1177/193229681200600632] [Medline: 23294798]
10. The internet pharmacy market in 2016. LegitScript for Center for Safe Internet Pharmacies. 2016. URL: https://safemedsonline.org/wp-content/uploads/2016/01/The-Internet-Pharmacy-Market-in-2016.pdf [accessed 2021-12-20]
11. Mackey TK, Nayyar G. Digital danger: a review of the global public health, patient safety and cybersecurity threats posed by illicit online pharmacies. Br Med Bull 2016 Jun;118(1):110-126 [FREE Full text] [doi: 10.1093/bmb/ldw016] [Medline: 27151957]
12. WHO Global Observatory for eHealth. Safety and Security on the Internet: Challenges and Advances in Member States : Based on the Findings of the Second Global Survey on eHealth. Geneva: World Health Organization; 2011:1-86.
13. Digital pharmacy accreditation criteria. National Association of Boards of Pharmacy. URL: https://nabp.pharmacy/programs/digital-pharmacy/ [accessed 2021-12-20]
14. Veronin MA, Clancy KM. Attrition of Canadian Internet pharmacy websites: what are the implications? Drug Healthc Patient Saf 2013;5:171-179 [FREE Full text] [doi: 10.2147/DHPS.S46232] [Medline: 23983491]
15. Penley B, Chen H, Eckel SF, Ozawa S. Characteristics of online pharmacies selling Adderall. J Am Pharm Assoc (2003) 2021;61(1):103-109 [FREE Full text] [doi: 10.1016/j.japh.2020.07.022] [Medline: 32912756]
16. Alwon BM, Solomon G, Hussain F, Wright DJ. A detailed analysis of online pharmacy characteristics to inform safe usage by patients. Int J Clin Pharm 2015 Feb;37(1):148-158. [doi: 10.1007/s11096-014-0056-1] [Medline: 25564180]
17. Akinbosoye OE, Taitel MS, Grana J, Hill J, Wade RL. Improving medication adherence and health care outcomes in a commercial population through a community pharmacy. Popul Health Manag 2016 Dec;19(6):454-461 [FREE Full text] [doi: 10.1089/pop.2015.0176] [Medline: 27035728]
18. Cicero TJ, Ellis MS. Health outcomes in patients using no-prescription online pharmacies to purchase prescription drugs. J Med Internet Res 2012 Dec 06;14(6):e174 [FREE Full text] [doi: 10.2196/jmir.2236] [Medline: 23220405]
19. LegitScript merchant and website classification guide. LegitScript. URL: https://www.legitscript.com/classifications-guide/ [accessed 2021-12-20]
20. Official measure of the digital world. SimilarWeb. URL: https://www.similarweb.com/ [accessed 2021-12-20]
21. Identify geographical location and proxy by IP address. IP2Location. URL: https://www.ip2location.com/ [accessed 2021-12-20]
22. Orizio G, Rubinelli S, Schulz PJ, Domenighini S, Bressanelli M, Caimi L, et al. "Save 30% if you buy today". Online pharmacies and the enhancement of peripheral thinking in consumers. Pharmacoepidemiol Drug Saf 2010 Sep;19(9):970-976. [doi: 10.1002/pds.2007] [Medline: 20652863]
23. Gelatti U, Pedrazzani R, Marcantoni C, Mascaretti S, Repice C, Filippucci L, et al. 'You've got m@il: fluoxetine coming soon!': accessibility and quality of a prescription drug sold on the web. Int J Drug Policy 2013 Sep;24(5):392-401. [doi: 10.1016/j.drugpo.2013.01.006] [Medline: 23433969]
24. Prescription savings you can trust. PharmacyChecker. URL: https://www.pharmacychecker.com [accessed 2021-12-20]
25. HealthWarehouse. URL: https://www.healthwarehouse.com/ [accessed 2021-12-20]
26. CVS Pharmacy. URL: https://www.cvs.com/ [accessed 2021-12-20]
27. Prescription same-day home delivery on the Costco app. Costco Pharmacy. URL: https://costco.mobilepharmacyhelp.com/same-day-delivery [accessed 2021-12-20]
28. Findlay S. States pass record number of laws to reel in drug prices. Kaiser Health News. 2019. URL: https://khn.org/news/states-pass-record-number-of-laws-to-reel-in-drug-prices/ [accessed 2021-12-20]

29. Personal importation. US Food & Drug Administration. 2018. URL: https://www.fda.gov/industry/import-basics/personal-importation [accessed 2021-12-20]
30. Lilly to introduce lower-priced insulin. Eli Lilly and Company. 2019. URL: https://www.prnewswire.com/news-releases/lilly-to-introduce-lower-priced-insulin-300805560.html [accessed 2021-11-23]
31. Inaccessible insulin: the broken promise of Eli Lilly's authorized generic. Prepared by the Offices of U.S. Senator Elizabeth Warren and U.S. Senator Richard Blumenthal. 2019. URL: https://www.warren.senate.gov/imo/media/doc/Inaccessible%20Insulin%20report.pdf [accessed 2021-11-23]
32. President Trump announces lower out of pocket insulin costs for Medicare's seniors. Centers for Medicare & Medicaid Services. 2020. URL: https://www.cms.gov/newsroom/press-releases/president-trump-announces-lower-out-pocket-insulin-costs-medicares-seniors [accessed 2021-11-23]
33. Berchick E, Barnett J, Upton R. Health insurance coverage in the United States. In: Kinney ED, editor. Protecting American Health Care Consumers. Durham, North Carolina: Duke University Press; 2019.
34. Budnitz DS, Lovegrove MC, Shehab N, Richards CL. Emergency hospitalizations for adverse drug events in older Americans. N Engl J Med 2011 Nov 24;365(21):2002-2012. [doi: 10.1056/NEJMsa1103053] [Medline: 22111719]
35. Basu S, Berkowitz SA, Seligman H. The monthly cycle of hypoglycemia: an observational claims-based study of emergency room visits, hospital admissions, and costs in a commercially insured population. Med Care 2017 Jul;55(7):639-645 [FREE Full text] [doi: 10.1097/MLR.0000000000000728] [Medline: 28481762]
36. American Diabetes Association. 3. Comprehensive medical evaluation and assessment of comorbidities: standards of medical care in diabetes-2018. Diabetes Care 2018 Dec;41(Suppl 1):28-37. [doi: 10.2337/dc18-S003] [Medline: 29222374]
37. Ozawa S, Evans DR, Bessias S, Haynie DG, Yemeke TT, Laing SK, et al. Prevalence and estimated economic burden of substandard and falsified medicines in low- and middle-income countries: a systematic review and meta-analysis. JAMA Netw Open 2018 Aug 03;1(4):e181662 [FREE Full text] [doi: 10.1001/jamanetworkopen.2018.1662] [Medline: 30646106]
38. Orizio G, Merla A, Schulz PJ, Gelatti U. Quality of online pharmacies and websites selling prescription drugs: a systematic review. J Med Internet Res 2011 Sep 30;13(3):e74 [FREE Full text] [doi: 10.2196/jmir.1795] [Medline: 21965220]
39. BeSafeRx: Your source for online pharmacy information. US Food & Drug Administration. 2016. URL: https://www.fda.gov/drugs/quick-tips-buying-medicines-over-internet/besaferx-know-your-online-pharmacy [accessed 2021-12-20]
40. Alliance for safe online pharmacies. ASOP Global. URL: https://buysaferx.pharmacy/ [accessed 2021-12-20]
41. Verify before you buy. Center for Safe Internet Pharmacies (CSIP). URL: https://verifybeforeyoubuy.org/ [accessed 2021-12-20]
42. Schlosser A, Crowley T. North Carolina Secretary of State partners with CSIP to promote online pharmacy verification. The Center for Safe Internet Pharmacies (CSIP). URL: https://safemedsonline.org/2020/05/north-carolina-secretary-of-state-partners-with-csip-to-promote-online-pharmacy-verification/ [accessed 2021-12-20]
43. Operation Pangea - shining a light on pharmaceutical crime. Interpol. 2019. URL: https://www.interpol.int/en/News-and-Events/News/2019/Operation-Pangea-shining-a-light-on-pharmaceutical-crime [accessed 2021-12-20]
44. LegitScript helps shut down 6,149 rogue internet pharmacies in 2019. LegitScript. 2020. URL: https://www.legitscript.com/blog/2020/01/legitscript-helps-shut-down-6149-rogue-internet-pharmacies-in-2019/ [accessed 2021-12-20]
45. Charm T, Coggins B, Robinson K, Wilkie J. The great consumer shift: Ten charts that show how US shopping behavior is changing. McKinsey & Company 2020:1-13 [FREE Full text]

Abbreviations

FDA: Food and Drug Administration
NABP: National Association of Boards of Pharmacy

Edited by R Kukafka; submitted 18.11.20; peer-reviewed by S Amini, J Appalasamy; comments to author 15.12.20; revised version received 30.12.20; accepted 23.09.21; published 14.02.22.

Please cite as:
Penley B, Minshew L, Chen HH, Eckel S, Ozawa S
Accessibility of Low-cost Insulin From Illegitimate Internet Pharmacies: Cross-sectional Study
J Med Internet Res 2022;24(2):e25855
URL: https://www.jmir.org/2022/2/e25855
doi:10.2196/25855
PMID:35156937

Original Paper

COVID-19 Treatments Sold Online Without Prescription Requirements in the United States: Cross-sectional Study Evaluating Availability, Safety and Marketing of Medications

Sachiko Ozawa[1,2], MHS, PhD; Joanna Billings[1], PharmD; Yujiao Sun[1], PharmD; Sushan Yu[1], PharmD; Benjamin Penley[1], PharmD

[1]Division of Practice Advancement and Clinical Education, Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[2]Department of Maternal and Child Health, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

Corresponding Author:
Sachiko Ozawa, MHS, PhD
Division of Practice Advancement and Clinical Education
Eshelman School of Pharmacy
University of North Carolina at Chapel Hill
CB# 7574, Beard Hall 115G
Chapel Hill, NC, 27599
United States
Phone: 1 919 966 2626
Email: ozawa@unc.edu

Abstract

Background: The COVID-19 pandemic has increased online purchases and heightened interest in existing treatments. Dexamethasone, hydroxychloroquine, and lopinavir-ritonavir have been touted as potential COVID-19 treatments.

Objective: This study assessed the availability of 3 potential COVID-19 treatments online and evaluated the safety and marketing characteristics of websites selling these products during the pandemic.

Methods: A cross-sectional study was conducted in the months of June 2020 to August 2020, by searching the first 100 results on Google, Bing, and Yahoo! mimicking a US consumer. Unique websites were included if they sold targeted medicines, were in English, offered US shipping, and were free to access. Identified online pharmacies were categorized as rogue, unclassified, or legitimate based on LegitScript classifications. Patient safety characteristics, marketing techniques, price, legitimacy, IP addresses, and COVID-19 mentions were recorded.

Results: We found 117 websites: 30 selling dexamethasone (19/30, 63% rogue), 39 selling hydroxychloroquine (22/39, 56% rogue), and 48 selling lopinavir-ritonavir (33/48, 69% rogue). This included 89 unique online pharmacies: 70% were rogue (n=62), 22% were unapproved (n=20), and 8% were considered legitimate (n=7). Prescriptions were not required among 100% (19/19), 61% (20/33), and 50% (11/22) of rogue websites selling dexamethasone, lopinavir-ritonavir, and hydroxychloroquine, respectively. Overall, only 32% (24/74) of rogue websites required prescriptions to buy these medications compared with 94% (31/33) of unapproved and 100% (10/10) of legitimate websites (P<.001). Rogue sites rarely offered pharmacist counseling (1/33, 3% for lopinavir-ritonavir to 2/22, 9% for hydroxychloroquine). Drug warnings were unavailable in 86% (6/7) of unapproved dexamethasone sites. It was difficult to distinguish between rogue, unapproved, and legitimate online pharmacies solely based on website marketing characteristics. Illegitimate pharmacies were more likely to offer bulk discounts and claim price discounts, yet dexamethasone and hydroxychloroquine were more expensive online. An inexpensive generic version of lopinavir-ritonavir that is not authorized for use in the United States was available online offering US shipping. Some websites claimed hydroxychloroquine and lopinavir-ritonavir were effective COVID-19 treatments despite lack of scientific evidence. In comparing IP addresses to locations claimed on the websites, only 8.5% (7/82) matched their claimed locations.

Conclusions: The lack of safety measures by illegitimate online pharmacies endanger patients, facilitating access to medications without appropriate oversight by health care providers to monitor clinical response, drug interactions, and adverse effects. We demonstrated how easy it is to go online to buy medications that are touted to treat COVID-19 even when current clinical evidence does not support their use for self-treatment. We documented that illegitimate online pharmacies sidestep prescription requirements,

skirt pharmacist counseling, and make false claims regarding efficacy for COVID-19 treatment. Health care professionals must urgently educate the public of the dangers of purchasing drugs from illegitimate websites and highlight the importance of seeking treatment through authentic avenues of care.

(J Med Internet Res 2022;24(2):e27704) doi:10.2196/27704

KEYWORDS

COVID-19; medication; internet; online pharmacy; drug

Introduction

The coronavirus disease 2019 (COVID-19) pandemic has intensified the demand for effective medications and made online purchases commonplace. According to a recent global study of 5000 consumers, over one-third (36%) of consumers reported shopping online weekly since the pandemic, which is up from 28% pre-COVID-19 [1]. As online shopping becomes more common, the rate of online purchase of medications is also increasing [2]. Online pharmacies have grown from US $29.35 billion in 2014 to a projected global market size of US $128 billion by 2023 [3].

Yet, consumers may not be aware of how prevalent illicit pharmacies are online. In 2016, there were 30,000-35,000 online pharmacies accessible to US customers on the internet [4]. Of these pharmacies, 96% were found to be illegitimate and in violation of US pharmacy laws and practice standards [5]. Illegitimate online pharmacies pose a significant safety threat to consumers, as their products cannot be guaranteed; the manufacturing, storage, and shipping conditions of the drugs may not be regulated; and they bypass pharmacists who can ensure proper dosages and use of medications [6].

As the pandemic continues to claim many lives [7], there has been heightened interest in finding potential treatments for COVID-19 among existing medications. Several existing medications have been touted as potential treatments for COVID-19, including dexamethasone, hydroxychloroquine, and lopinavir-ritonavir [8]. Dexamethasone is a glucocorticoid agent used for several disease states for its anti-inflammatory and immunosuppressive properties. It has been trialed in patients with COVID-19 to see if it may reduce disease progression to respiratory failure and death [9]. A controlled, open-label trial among patients who were hospitalized with COVID-19 found that the use of dexamethasone resulted in lower 28-day mortality among those who were receiving either invasive mechanical ventilation or oxygen alone [9,10]. This trial, along with others, supports the idea that dexamethasone can prevent mortality in hospitalized patients with COVID-19 [11]. Such supportive evidence behind the effectiveness of dexamethasone has caused an increase in its use [12]. It is important to note though that corticosteroids such as dexamethasone are not benign drugs and should only be taken under the direct supervision of a medical provider. If misused, dexamethasone can cause side effects such as hypertension, endocrine abnormalities, vision problems, pancreatitis, and osteoporosis [13]. Monitoring the use of dexamethasone is therefore required to assess side effects and response to therapy.

Hydroxychloroquine has also been considered a potential COVID-19 treatment and prophylaxis [14]. Hydroxychloroquine is an aminoquinoline anti-infective agent indicated for the treatment of malaria. Despite promising mechanistic properties, several randomized clinical trials found that hydroxychloroquine did not prevent illness or confirmed infection when used as postexposure prophylaxis within 4 days after moderate or high-risk exposure to COVID-19 [14,15]. However, endorsement of the medication by high-profile figures including former US President Donald Trump during early phases of the pandemic drew public attention to the medication. It was reported that Google search volume for purchasing hydroxychloroquine increased by 14-fold in March 2020, when the clinical efficacy for the medication was still inconclusive [16]. The use of hydroxychloroquine for COVID-19 can be problematic especially without a health care prescription, putting patients at risk for insufficient laboratory monitoring and side effects such as nausea, low blood sugar, movement disorders, eye problems, and cardiac problems [10].

Another medication that has been touted for COVID-19 treatment is lopinavir-ritonavir. Lopinavir-ritonavir is an antiretroviral agent that is used for HIV infection suppression. Since lopinavir had in vitro inhibitory activity against the SARS-CoV virus in 2003 and the SARS-CoV-2 (COVID-19) virus is genetically similar, clinical trials were conducted in patients with COVID-19 [17]. However, a randomized, controlled, open-label trial involving hospitalized adult patients with severe COVID-19 found no benefit of lopinavir-ritonavir treatment beyond standard care [14]. Further studies and case reports have reinforced these findings, and lopinavir-ritonavir is generally not recommended for treatment of COVID-19 [18,19]. If consumers use lopinavir-ritonavir in efforts to combat a COVID-19 infection, they would not only be using an ineffective treatment but would also be exposing themselves to unnecessary side effects.

We sought to assess the accessibility of dexamethasone, hydroxychloroquine, and lopinavir-ritonavir online during the COVID-19 pandemic and to analyze their implications for medication safety. In doing so, we recorded medication prices, marketing practices, and statements related to COVID-19 from websites selling these medications. Our aim was to assess what US consumers would find if they utilized online pharmacies looking for medications to treat COVID-19 amid the pressures of a global pandemic.

Methods

We searched for online pharmacy websites selling dexamethasone, hydroxychloroquine, or lopinavir-ritonavir

using 3 search engines (Google, Bing, and Yahoo!) between June 2020 and August 2020. Two searches were conducted on each search engine using the brand and generic names of each medication. We used the search terms "buy [drug name] online" (eg, "buy lopinavir-ritonavir online" and "buy Kaletra online") to simulate patients seeking to buy these treatments online. We screened the first 100 search results for every search term and recorded the results of our searches. Websites were included in this analysis if they were published in the English language, were free to access, were active, claimed to sell the drugs of interest, had a unique URL, and were classified by LegitScript. Websites were excluded if they did not sell the medications of interest or did not ship to the United States. Screenshots of the web pages were taken at the time the pages were accessed.

The legitimacy of included websites was assessed by LegitScript, an online service that monitors online pharmacies for compliance with applicable laws and regulations. LegitScript classifies online pharmacies based on licensure or registration in appropriate jurisdictions, sales of controlled substances, history of disciplinary sanctions, requirement of valid prescription, compliance with applicable laws, protection of privacy, patient services offered, transparency, and domain name registration [4]. This analysis utilized LegitScript's classification, where *rogue* pharmacies "engage in illegal, unsafe, or misleading activities like selling prescription drugs without a prescription," *unapproved* pharmacies face "some problem with regulatory compliance or risk, such as operating legally in one jurisdiction but not in others," and *legitimate pharmacies* "passed LegitScript's certification criteria" [4]. Rogue and unapproved pharmacies were considered illegitimate pharmacies in this analysis.

In order to understand how online pharmacies attend to patient safety, we assessed whether prescriptions were required to purchase these medications or if patients needed to fill out a health-related questionnaire prior to ordering medications. We assessed whether pharmacist counseling was available, appropriate drug information was on the drug product page, warnings in using the medications were noted, and there was a quantity limit on purchases. Quantity control was defined as the website setting a maximum allowable quantity (or day supply) of medication that was available for purchase on the checkout page. Claims made related to the efficacy of these medications in treating COVID-19 were also abstracted from the websites. Patient safety analyses were conducted for each drug individually. Chi-squared tests were conducted to examine differences in patient safety characteristics by pharmacy legitimacy and across medications.

We recorded some marketing characteristics of websites, including whether the website claimed a price discount compared with brick-and-mortar pharmacies, availability of bulk product price discounts, and offers of coupons or promotion codes. Marketing characteristics of rogue, unapproved, and legitimate sites were compared using chi-squared tests. We tracked whether online pharmacies offered phone numbers or WhatsApp accounts and whether consumers could speak with sales associates. Websites that claimed to be registered with pharmacy-governing organizations were recorded, as well as customer testimonies and privacy assurances. Marketing of COVID-19–related products such as ads for masks, sanitizing wipes, and hand sanitizers, as well as explicit mentions of COVID-19 on websites, were also recorded. Quotes about COVID-19 from these websites were classified by website legitimacy and were analyzed qualitatively, examining themes that emerged from texts. Unique websites were defined as those with unique URLs, and marketing analyses were conducted once per unique website (ie, if a website sold 2 of our medications of interest, we reported the website's marketing characteristic once).

Prices and shipping costs for obtaining 60 dexamethasone 0.5 mg tablets, hydroxychloroquine 200 mg capsules, and lopinavir-ritonavir 200 mg-50 mg tablets were collected from each website. These quantity and strengths were chosen as they were the most commonly available for patients to select and purchase presented on the web pages. Online pharmacy prices at checkout were compared with prices offered through GoodRx, which tracks prescription drug prices in the United States. Where possible, we compared prices of generic versions of the medications. However, the GoodRx price for lopinavir-ritonavir reflected the price for the brand name drug Kaletra in the United States, as a generic version had not been approved for use in the United States. Country locations reported on pharmacy websites were also recorded and compared with registered locations of the IP addresses identified via IP2location [20]. Where available, we assessed the monthly traffic volume of pharmacy sites using SimilarWeb, a site that specializes in web analytics and reports traffic volumes from open exchanges of first-party data, and by surveying public data sources [21].

Results

Across 3 search engines, we identified a total of 117 websites that claimed to sell dexamethasone, hydroxychloroquine, or lopinavir-ritonavir online. This included 30 websites selling dexamethasone, 39 sites selling hydroxychloroquine, and 48 sites selling lopinavir-ritonavir. The majority of websites were illegitimate—we only found 4 (4/30, 13%), 5 (5/39, 13%), and 1 (1/48, 2%) legitimate websites selling dexamethasone, hydroxychloroquine, and lopinavir-ritonavir, respectively. We found that 63% (19/30) of websites selling dexamethasone, 56% (22/39) of sites selling hydroxychloroquine, and 69% (33/48) of sites selling lopinavir-ritonavir were rogue based on LegitScript classifications. An additional 26 websites that were not classified by LegitScript were excluded from our analysis.

Rogue websites tended to be distinct from unapproved (31/33, 94%) and legitimate sites (10/10, 100%), with only 32% (24/74) of sites requiring prescriptions to buy these medications (P<.001; Table 1). Health-related questionnaires were seldom used across all websites (utilized by 1 rogue and 7 unapproved websites). Overall, few (15/117, 12.8%) websites made overt offers for consumers to speak with a pharmacist, including legitimate sites (1/10, 10%). Medication warnings were often missing from unapproved websites (13/33, 39%; P<.001). Quantity control was rarely seen in legitimate sites, as prescriptions were meant to limit purchasable quantity.

Table 1. Patient safety characteristics stratified by rogue, unapproved, and legitimate online pharmacies.

Patient safety characteristics	Rogue (n=74), n (%)	Unapproved (n=33), n (%)	Legitimate (n=10), n (%)	*P* value
Prescription required	24 (32)	31 (94)	10 (100)	<.001
Health-related questionnaire	1 (1)	7 (21)	0 (0)	<.001
Offer to speak with pharmacist	4 (5)	10 (30)	1 (10)	.002
Drug information on product page	44 (59)	18 (55)	9 (90)	.12
Drug related warnings on product page	60 (81)	13 (39)	8 (80)	<.001
Quantity control	45 (61)	27 (82)	2 (20)	.001

Across medications, many illegitimate websites did not require prescriptions to obtain dexamethasone (21/26, 81%), compared with hydroxychloroquine (11/34, 32%) and lopinavir-ritonavir (20/47, 43%; *P*=.02; Table 2). Rogue websites by far had the most lenient prescription requirements, with none (0/19, 0%) of the rogue websites selling dexamethasone requiring prescriptions, while 50% (11/22) and 39% (13/33) of rogue websites selling hydroxychloroquine and lopinavir/ritonavir, respectively, required prescriptions. Medication information and warnings were particularly lacking for illegitimate websites selling hydroxychloroquine and lopinavir-ritonavir, with only 65% (22/34) and 36% (17/47) of websites, respectively, providing accurate drug information. Among illegitimate websites, 53% (18/34) selling hydroxychloroquine and 77% (36/47) selling lopinavir-ritonavir provided medication-related warnings. Across all websites selling dexamethasone, 87% (26/30) included medicine information, and 73% (22/30) provided medication-related warnings on their web pages.

Table 2. Patient safety characteristics stratified by online pharmacy legitimacy and medication.

Patient safety characteristics	Rogue, n (%)			Unapproved, n (%)			Legitimate, n (%)			*P* value
	DEX[a] (n=19)	HCQ[b] (n=22)	L-R[c] (n=33)	DEX (n=7)	HCQ (n=12)	L-R (n=14)	DEX (n=4)	HCQ (n=5)	L-R (n=1)	
Prescription required(yes)	0 (0)	11 (50)	13 (39)	5 (71)	12 (100)	14 (100)	4 (100)	5 (100)	1 (100)	.02
Health-related questionnaire (yes)	0 (0)	1 (5)	0 (0)	0 (0)	3 (25)	4 (29)	0 (0)	0 (0)	0 (0)	.29
Offer to speak with pharmacist(yes)	1 (5)	2 (9)	1 (3)	1 (14)	6 (50)	3 (21)	0 (0)	1 (20)	0 (0)	.87
Drug information on product page(yes)	18 (95)	14 (64)	12 (36)	5 (71)	8 (67)	5 (36)	3 (75)	5 (100)	1 (100)	.59
Drug-related warnings on product page (yes)	18 (95)	12 (54)	30 (91)	1 (14)	6 (50)	6 (43)	3 (75)	4 (80)	1 (100)	.06
Quantity control (yes)	15 (79)	11 (50)	19 (58)	7 (100)	12 (100)	8 (57)	1 (25)	1 (20)	0 (0)	.37

[a]DEX: dexamethasone.

[b]HCQ: hydroxychloroquine.

[c]L-R: lopinavir-ritonavir.

Marketing characteristics were examined across unique websites. We identified 89 unique websites that were classified by LegitScript, as 28 websites sold more than one product of interest. Among the 89 websites, 62 (70%) were identified as rogue, 20 (22%) were unapproved, and only 7 (8%) websites were legitimate.

When examining marketing characteristics of websites, it was more difficult to distinguish between rogue, unapproved, and legitimate online pharmacies (Table 3). Privacy assurances were most frequently offered by unapproved pharmacies (17/20, 85%), compared with 52% (32/62) for rogue and 57% (4/7) for legitimate pharmacies (*P*=.03). Rogue and unapproved websites appeared to be more likely to claim price discounts (58/62, 94% and 18/20, 90% respectively), while 71% (5/7) of legitimate websites made a similar claim (*P*=.12). Rogue and unapproved websites tended to offer more bulk discounts (41/62, 66% and 13/20, 65%, respectively) compared with legitimate websites (2/7, 29%; *P*=.15). It is notable that 32% (20/62) of rogue websites and 60% (12/20) of unapproved websites claimed to be registered by pharmacy-governing organizations. The most common registration claims made by illegitimate websites included the Canadian International Pharmacy Association (CIPARx; 33/82, 40%) and PharmacyChecker (25/82, 30%). Legitimate websites mostly claimed registration with the National Association of Boards of Pharmacy (2/7, 29%), the Better Business Bureau (4/7, 57%), or LegitScript (5/7, 71%).

Table 3. Marketing characteristics stratified by rogue, unapproved, and legitimate online pharmacies (n=89).

Marketing characteristics	Rogue (n=62), n (%)	Unapproved (n=20), n (%)	Legitimate (n=7), n (%)	*P* value[a]
Claims price discount (yes)	58 (94)	18 (90)	5 (71)	.12
Bulk discounts (yes)	41 (66)	13 (65)	2 (29)	.15
Coupon or promo code (yes)	37 (60)	8 (40)	4 (57)	.29
Phone # or WhatsApp (yes)	56 (90)	19 (95)	7 (100)	.56
Offer to speak with an associate (yes)	45 (72)	16 (80)	6 (86)	.60
Registration claims (yes)	20 (32)	12 (60)	3 (43)	.08
Customer testimonials (yes)	39 (63)	8 (40)	5 (71)	.15
Privacy claims (yes)	32 (52)	17 (85)	4 (57)	.03

[a]Analysis was conducted across websites selling dexamethasone, hydroxychloroquine, and/or lopinavir-ritonavir, which were classified by LegitScript for legitimacy.

A total of 7 illegitimate websites made unsubstantiated claims regarding the efficacy of hydroxychloroquine (n=4) or lopinavir-ritonavir (n=3) in treating COVID-19 (Table 4). Illegitimate pharmacies made a wide variety of uncorroborated medication efficacy claims from “hydroxychloroquine has shown positive results when being used to treat patients with COVID-19” to “[lopinavir-ritonavir] has been tested and included in the protocols of treatment of the novel coronavirus COVID-19.” One illegitimate website quoted former US President Donald Trump’s Twitter account to note that hydroxychloroquine should be used to treat COVID-19. Four illegitimate websites created a COVID-19 medication category and tagged hydroxychloroquine or lopinavir-ritonavir. Illegitimate websites also noted supply limitations: “[Hydroxychloroquine is] not available at this time due to COVID-19 shipping disruptions.” COVID-19–related supplies such as gloves, masks, or test kits were advertised by 4 illegitimate websites.

On the other hand, legitimate pharmacy websites rarely mentioned COVID-19 on their web pages, neither claiming nor denying the efficacy of these medications related to COVID-19. Two legitimate websites mentioned limitations in dispensing hydroxychloroquine due to COVID-19, limiting supplies to existing hydroxychloroquine patients or restricting hydroxychloroquine to only patients with a positive COVID-19 test. One legitimate website advertised hand sanitizers and masks alongside these medications.

Prices, with and without standard shipping, for 60 tablets of each medication bought online (0.5 mg tablets of dexamethasone, 200 mg capsules of hydroxychloroquine, and 200 mg lopinavir and 50 mg ritonavir tablets) were compared with the equivalent quantity price at brick-and-mortar pharmacies listed on GoodRx (Figure 1). Dexamethasone and hydroxychloroquine were both more expensive online, with or without shipping, compared with listed average GoodRx prices. Average prices were lower at brick-and-mortar stores for 0.5 mg tablets of dexamethasone at US $10.74, compared with US $38.07 for the medication online and US $50.54 with shipping. Similarly, 200 mg capsules of hydroxychloroquine were cheaper at brick-and-mortar stores at US $39.64 on average, compared with US $83.22 online and US $96.04 with shipping. A generic version of lopinavir-ritonavir was available online for US consumers and sold on average at US $211.77 (US $229.52 with shipping), even though this generic drug has not been approved for use in the United States. This unapproved generic lopinavir-ritonavir was much cheaper online than average GoodRx prices for brand-name drug Kaletra at US $532.40.

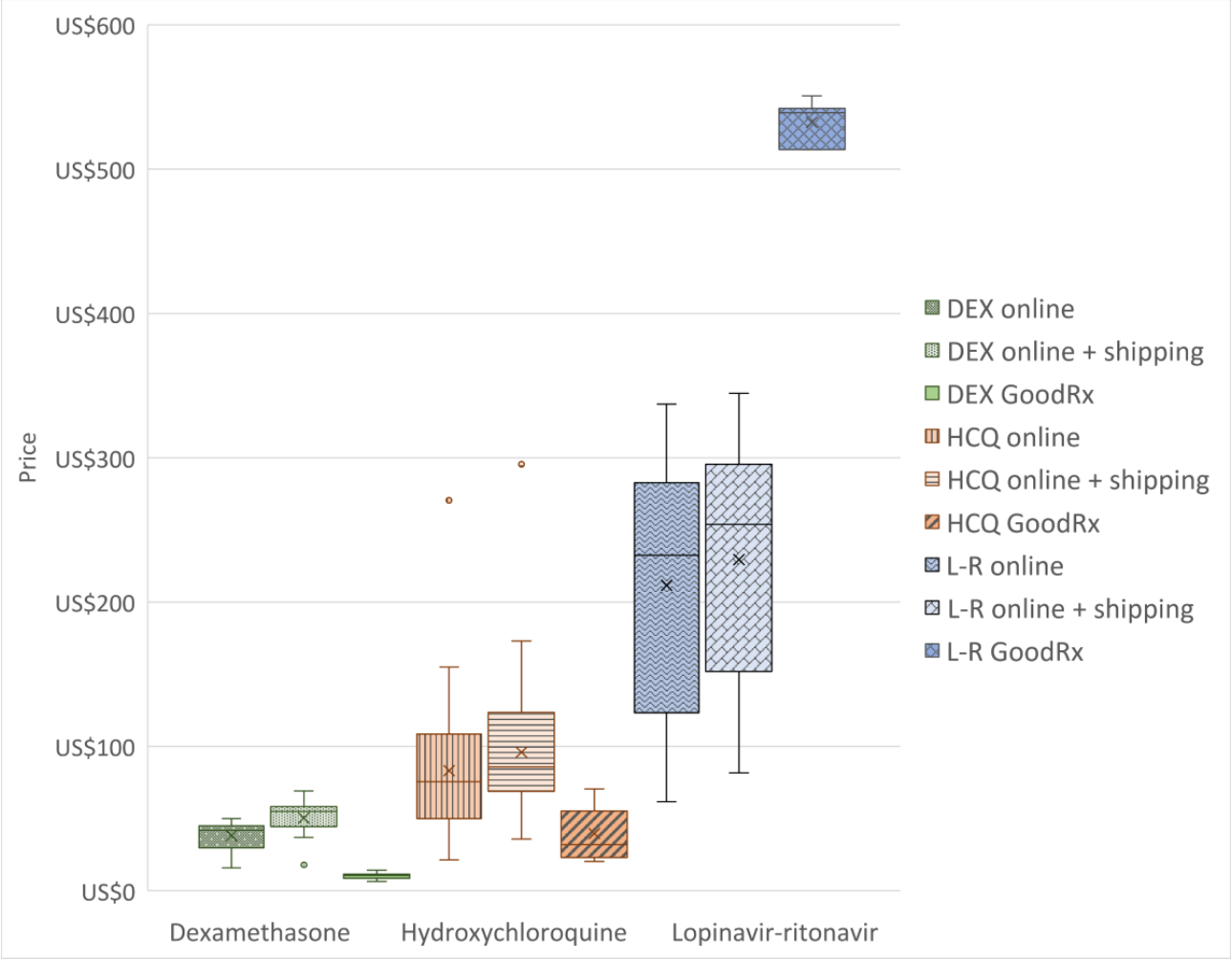
Table 4. COVID-19 mentions on illegitimate and legitimate online pharmacy websites.

Type of COVID-19 mentions	Quotes	Mentions[a]
Quotes from illegitimate pharmacies		
Direct claims that medication may be effective for COVID-19	• "drug [hydroxychloroquine] confirmed effective on COVID-19" • "It was announced by American president that Hydroxychlorquine should be used to treat coronavirus, Donald Trump said this in his Twitter on March 21 2020." • "[study] proved that Hydroxychloroquine… can cut the death rate significantly in sick, hospitalized COVID patients without heart-related side effects... The analysis of the study shows hydroxychloroquine helped save lives." • "Hydroxychloroquinine has shown positive results when being used to treat patients with Covid-19 (coronavirus)." • "it was found out that Kaletra shows positive results in a blockage of a COVID-19 viral replication" • "This HIV medication [lopinavir-ritonavir] is one of the few drugs effective for COVID-19 treatment)" • "It [lopinavir-ritanovir] has been tested and included in the protocols of treatment of the novel coronavirus COVID-19"	7
Claims medication may be effective for COVID-19	• "Reports have shown that this medicine [hydroxychloroquine] might be effective against Coronavirus, but it has not been proven" • "Reports have shown that this medicine [lopinavir-ritonavir] might be effective against Coronavirus, but it has not been proven."	2
Supply/shipping limitation due to COVID-19	• "…New Zealand, in an effort to preserve adequate stocks of this medicine for their own people during the Covid-19 pandemic have halted all exports until further notice" • "Limited supply. Currently being dispensed to existing Plaquenil patients only" • "Not available at this time due to COVID-19 shipping disruptions."	3
Precaution regarding efficacy in COVID-19	• "Taking this medication will not prevent you from passing HIV or COVID-19 to other people"	1
Medication included in COVID-19 medication category	• Tag[b]: COVID-19 [hydroxychloroquine] • Tag[b]: coronavirus [lopinavir-ritanovir] [hydroxychloroquine] • Tag[b]: covid19 treatment, covid19 medications uk [lopinavir-ritanovir]	4
Ads for products related to COVID-19	• Gloves • COVID test kits • Masks	4
Quotes from legitimate pharmacies		
Indication limitation or dispense restriction due to COVID-19	• "we are ONLY dispensing Hydroxychloroquine to patients with history of use for an autoimmune disorder … prior to 3/1/2020 or a positive diagnosis of COVID-19… Doctors prescribing outside of their scope of practice will be DENIED" • "Due to the national shortage, we can only dispense this drug [hydroxychloroquine] if you have tested positive for coronavirus/COVID-19."	2
Ads for products related to COVID-19	• Sanitizer and masks	1

[a]Denotes the number of online pharmacy websites per type of COVID-19 mention.

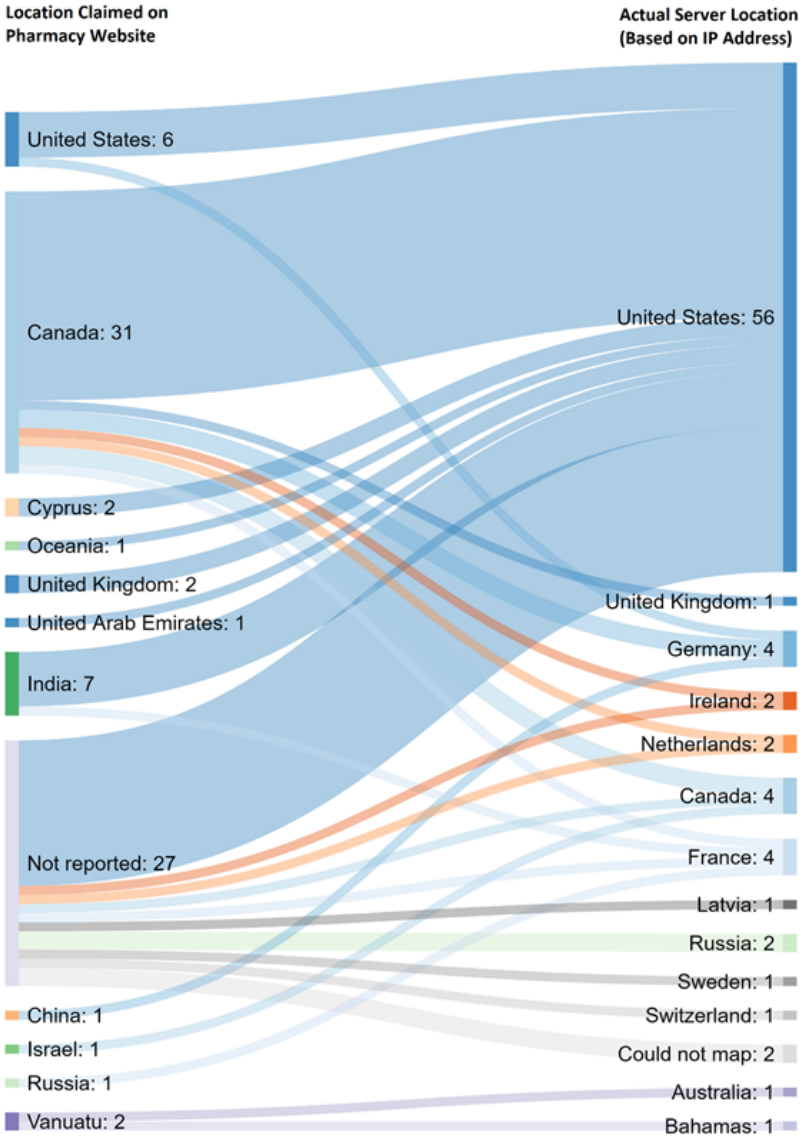
[b]These websites created COVID-19 medication categories or categorized the medication under a specific tag on the product pages.

Figure 1. Prices to purchase 60 tablets/capsules of dexamethasone (DEX), hydroxychloroquine (HCQ), and lopinavir-ritonavir (L-R) online versus GoodRx.



Large numbers of countries were involved in hosting online pharmacies, either where online pharmacies claimed to be or where their servers were located (Figure 2). Most of the illegitimate websites (75/82, 92%) hosted their servers in a country different from that listed on their website. The majority of illegitimate websites (58/82, 71%) either claimed to be a Canadian pharmacy (31/82, 38%) or did not list any country location (27/82, 33%) on their websites. Of the websites claiming to be located in Canada or India (38/82, 46%), the majority actually had IP addresses with a registered server in the United States (29/38, 76%). All legitimate websites (n=7) claimed to be in the United States and had IP addresses that indicated US servers. Website traffic data were available on SimilarWeb for 9 websites (3 rogue and 6 legitimate sites). Average website traffic for legitimate websites ranged from 10,200 to 10,900,000 monthly views, while rogue websites ranged from 10,000 to 20,000 monthly views.

Figure 2. Reported versus server locations of illegitimate online pharmacies (n=82).



Discussion

Principal Findings

We found that illegitimate pharmacy websites claiming to sell dexamethasone, hydroxychloroquine, and lopinavir-ritonavir endanger patient safety by bypassing prescription requirements, denying pharmacist counseling, and making false claims regarding efficacy in treating COVID-19 infections. We demonstrated that these websites are easy to find using common search engines and that they attempt to garner attention from potential consumers by claiming discounted prices and offering bulk discounts. We discuss in the following sections the reasons why these websites are harmful (lack of prescriptions, no pharmacist contact, and use of these medications for COVID-19 treatment) and how they try to lure consumers (medicine prices, online pharmacy marketing, and global transactions) [22].

Lack of Prescriptions

Many illegitimate websites we found selling dexamethasone (21/26, 81%), hydroxychloroquine (11/34, 32%), and lopinavir-ritonavir (20/47, 43%) did not require prescriptions. This is a recipe for tragedy as prescriptions ensure that health care providers have reviewed patients' medical histories and that selected medications are safe and appropriate. The lack of prescriptions needed to access these medications through online pharmacies is an immense patient safety concern, as these medications require intensive monitoring for side effects and therapeutic effects when they are initiated.

For example, dexamethasone requires a physician-guided dosing taper for safe and effective use. If dexamethasone is stopped abruptly after long-term use, it can cause hypothalamic-pituitary adrenal axis suppression, which can lead to a decreased immune response against infection or trauma [13]. Moreover, side effects of taking dexamethasone include hypertension, endocrine abnormalities, vision problems, pancreatitis, and osteoporosis [13]. For hydroxychloroquine, there have been reports of fatal cardiomyopathy and arrythmias with its use, which have led to recommendations for intensive cardiac monitoring during therapy initiation [23]. Hydroxychloroquine also has the potential to cause adverse effects to the liver, kidney, cerebellar cortex, pancreas, and eyes [23,24]. Lopinavir-ritonavir requires provider guidance for safe dosing and administration. An

overdose of lopinavir-ritonavir can result in cardiac toxicity, renal failure, and central nervous system depression [25].

No Pharmacist Contact

We also found that very few websites (15/117, 12.8%) offered pharmacist counseling before consumers purchase these medications, including even legitimate sites (1/10, 10%). Pharmacists ensure medication appropriateness and safety by ensuring that patients have the opportunity to clarify any dosing instructions, interactions, or side effects before they take medications [26]. With no pharmacist oversight, patients are more likely to experience drug interactions, incorrect dosing, or side effects that could have been avoided with counseling. For example, previous studies have examined the association between online pharmacy use, medication use, and health outcomes, where patients who obtained medications online were more likely to use medications at higher doses and more frequently and experience adverse events compared with patients obtaining medications from brick-and-mortar stores [27]. Additionally, most states in the United States have regulations in place where an offer to counsel with a pharmacist must be made at the time of medication dispensing [28]. Buying medications through online pharmacies bypasses these important patient safety measures and regulations.

Some adverse events specific to these medications could be avoided if pharmacists are able to review patients' profiles to analyze for interacting medications and counsel patients on potential side effects. For example, patients taking dexamethasone should not take nonsteroidal anti-inflammatory drugs as using both increases the risk of gastrointestinal side effects [13]. Since dexamethasone is metabolized in the liver, other medications that affect hepatic enzymes can also affect the metabolism of dexamethasone [13]. Likewise, hydroxychloroquine should also be used with caution in patients taking hepatotoxic medications or with hepatic disease [23]. Since hydroxychloroquine and lopinavir-ritonavir can increase the risk for cardiac arrythmia, it is important to avoid other medications that have a similar cardiac effect where possible [23,25]. Moreover, lopinavir-ritonavir interacts with any medications that induce, inhibit, or are metabolized by CYP3A, a very common metabolic liver enzyme [25]. If taken with a medication that is primarily cleared by CYP3A, lopinavir-ritonavir can cause buildup of the medication, which can lead to life-threatening side effects [25]. These examples demonstrate the importance for all online pharmacies to facilitate consumer interactions with pharmacists.

Use of These Medications for COVID-19 Treatment

Although some online pharmacy sites have touted the efficacy of these medications in treating COVID-19, current clinical evidence does not support their use for self-treatment of COVID-19 infections. Dexamethasone was found to be effective in reducing mortality in patients with COVID-19 in the RECOVERY trial; however, the benefit was only seen among patients with severe disease requiring supplemental oxygen or invasive mechanical ventilation [9]. Online purchasing of dexamethasone for self-treatment of COVID-19 is therefore not appropriate.

One of the most common references to medicine efficacy for COVID-19 treatment on pharmacy websites was related to hydroxychloroquine (n=14). Four illegitimate websites claimed hydroxychloroquine as an effective treatment for COVID-19, with 1 site citing former US President Trump's Twitter account as the source for this claim rather than scientific sources. Other sites created a COVID-19 medication category and deceptively included hydroxychloroquine in the category page. Hydroxychloroquine was one of the most studied medications during the early phase of the pandemic, with more than 50 registered clinical trials by May 2020 [29,30]. On April 27, 2020, the Food and Drug Administration (FDA) issued an emergency use authorization (EUA) for the use of hydroxychloroquine to treat adults and adolescents hospitalized with COVID-19 who are not able to participate in clinical trials. However, this EUA was revoked on June 15, 2020 after the FDA reviewed clinical evidence [31]. The FDA concluded that hydroxychloroquine is unlikely to be effective and the potential benefit of using the treatment no longer outweighs the risks of serious cardiac adverse events and other side effects associated with the medication [31]. The RECOVERY trial in the United Kingdom found no mortality benefit with hydroxychloroquine, while the Solidarity Trial coordinated by the World Health Organization terminated the hydroxychloroquine arm early based on interim analysis results that showed little or no reduction of mortality in hospitalized patients with COVID-19 [18,32].

We found one illegitimate website falsely claiming that lopinavir-ritonavir showed positive results in blocking COVID-19 viral replication. Such false statements can be hazardous for patients seeking reliable information. Similar to hydroxychloroquine, lopinavir-ritonavir was a popular explorative treatment at first, as a systematic review in 2020 including 9152 hospitalized patients found that it was the most frequently administered treatment, received by 21.9% of patients [33]. However, the RECOVERY trial also found the medication to not be effective, and the Solidarity Trial discontinued its lopinavir-ritonavir arm based on an interim analysis showing similar results [18,19].

Medicine Prices

Prior studies have suggested that the cost of medications online vis-à-vis brick-and-mortar stores differs depending on the product [34-37]. Here, we report similar findings. While the majority (76/82, 93%) of illegitimate online pharmacies claimed price discounts on their products, it was cheaper on average to purchase dexamethasone and hydroxychloroquine at brick-and-mortar pharmacies. We found that average online prices (excluding shipping) for dexamethasone and hydroxychloroquine were 285% and 110% more expensive, respectively, than GoodRx.

On the other hand, prices for lopinavir-ritonavir were cheaper on online pharmacies. In the United States, lopinavir-ritonavir 200 mg-50 mg tablets are only available as a brand name product, Kaletra, at an average GoodRx price of US $532.40. Meanwhile, illegitimate online pharmacies advertised a generic lopinavir-ritonavir that has not been approved in the United States at an average price of US $211.77. The availability of an

unapproved and cheaper "generic" lopinavir-ritonavir through online pharmacies means US consumers could in fact obtain a different medicine, a generic drug not approved in the United States, with significant safety concerns.

Online Pharmacy Marketing

Illegitimate websites were more likely to claim price discounts (76/82, 93%) than legitimate websites (5/7, 71%). They were also more likely to offer bulk discounts if the customer placed larger orders. Many illegitimate websites also made claims to be registered by pharmacy-governing organizations, claiming to be approved by various sites as a legitimate business practice. These registration claims often cited CIPARx or PharmacyChecker. It is important to note that LegitScript has found that CIPARx and PharmacyChecker have previously given approval to websites legally found to be illegitimate [17]. The lack of reliability of such accreditation sites are problematic and need to be corrected as less savvy customers may see a symbol of accreditation and believe the website to be legitimate.

Global Transactions

Most of the illegitimate online pharmacy websites analyzed in this study (75/82, 92%) hosted their servers in a country different from that listed on their website. This is in line with other studies that have reported similar differences in server and listed locations [35]. Although this difference does not verify that online pharmacies' physical locations are different from what are reported, it does raise the concern that many of these pharmacies are misleading patients by reporting to be located in one country but are actually dispensing from another country. The discrepancies also increase barriers to enforcement agencies' regulations. In addition, the pharmacies may be obtaining their medications from international manufacturers and suppliers that do not comply with US FDA regulations and good manufacturing practices. Worse yet, the medications dispensed by these illegitimate online pharmacies may be substandard or falsified and could cause harm to patients [38]. As 1 in 10 essential medications sold in low- and middle-income countries have been found to be substandard or falsified, some medications sold through online pharmacies could be substandard or falsified [39,40]. Global transactions facilitated by online pharmacies can be an especially opportunistic market during pandemics, when demand for COVID-19 treatments is high and regulatory actions may be inadequate [41]. Interventions are needed globally to ensure access to safe, quality-assured, and effective medications during the COVID-19 pandemic and beyond, with protections to minimize medication harm.

There is a number of limitations of our study to note. First, we did not attempt to capture the entirety of the online pharmacy landscape and were only able to present a snapshot of results at a point in time. We utilized 3 most commonly used search engines, conducted searches from the United States, conducted searches from the United States, and included unique sites from the first 10 pages of search results, imitating the behavior of a typical US consumer who may try to find online pharmacy websites selling these medications. We present a cross-sectional study where our data collection occurred in the summer of 2020. There may have been temporal changes to the availability of these medications and alterations in marketing tactics employed by online pharmacies since then. Second, this study focused solely on dexamethasone, hydroxychloroquine, and lopinavir-ritonavir, when the demand and supply for other medications to treat COVID-19 may have also been affected during this pandemic. Our medications of study were chosen based on their high publicity as a potential COVID-19 treatment during the first half of 2020.

We were also unable to uncover the actual purchasing frequency on these sites. We note that the web traffic to illegitimate online pharmacies may not be reflective of actual transaction volume, as many sites regularly close and open with new URLs or redirect links to avoid detection and regulation. However, the results we found demonstrate a snapshot of what an average US consumer may view while purchasing. Moreover, due to ethical concerns regarding providing fake prescriptions to these websites, we did not actually purchase medications from these online pharmacies. Therefore, we were unable to investigate the actual dispensing locations or test the quality of medications. Without following through with the full purchasing process, we could not confirm if pharmacist counseling is indeed offered after checkout. However, despite these limitations, we followed the systemic approach outlined in the methodology to assess the availability of high-profile medications during the COVID-19 pandemic from online pharmacies.

Conclusion

This analysis illustrates how easy it is to go online to buy medications that are touted to treat COVID-19 even when current clinical evidence does not support their use for self-treatment. Illegitimate online pharmacies endanger patient safety by sidestepping prescription requirements, skirting pharmacist counseling, and making false claims regarding efficacy in treating COVID-19 infections. It is important for health care providers to recognize the ease of access to these medications and raise awareness about current clinical evidence especially during the pandemic. Health care providers should also educate the public about the risks of purchasing medications from illegitimate websites, highlighting the importance of prescriptions and appropriate monitoring of therapeutic response, drug interactions, and adverse effects by pharmacists [42]. The amount of illegitimate online pharmacies that surfaced with our search also suggests a need for better search algorithms and stricter monitoring and regulations to prevent illegitimate pharmacies from being easily accessible online.

Acknowledgments

We thank the Chapel Hill Region Advanced Immersion program by the Eshelman School of Pharmacy at the University of North Carolina at Chapel Hill for facilitating this research project.

Conflicts of Interest

None declared.

References

1. Evans K. More than one-third of consumers shop online weekly since coronavirus hit. Digital Commerce 360. 2020 Oct 1. URL: https://www.digitalcommerce360.com/2020/10/01/more-than-one-third-of-consumers-shop-online-weekly-since-coronavirus-hit/ [accessed 2021-10-30]
2. Sabanoglu T. Share of U.S. consumers using online shopping before and after COVID-19 as of September 2020 and February 2021, by selected categories. Statista. URL: https://www.statista.com/statistics/1134709/consumers-us-online-purchase-before-after-covid-categories/ [accessed 2021-10-30]
3. Zhao H, Muthupandi S, Kumara S. Managing illicit online pharmacies: web analytics and predictive models study. J Med Internet Res 2020 Aug 25;22(8):e17239 [FREE Full text] [doi: 10.2196/17239] [Medline: 32840485]
4. About LegitScript. LegitScript. URL: https://www.legitscript.com/about/ [accessed 2021-10-30]
5. LegitScript.com. The Internet Pharmacy Market in 2016: Trends, Challenges, and Opportunities. The Center for Safe Internet Pharmacies (CSIP). 2016 Jan. URL: https://safemedsonline.org/wp-content/uploads/2016/01/The-Internet-Pharmacy-Market-in-2016.pdf [accessed 2021-10-30]
6. Fittler A, Vida RG, Káplár M, Botz L. Consumers turning to the internet pharmacy market: cross-sectional study on the frequency and attitudes of Hungarian patients purchasing medications online. J Med Internet Res 2018 Aug 22;20(8):e11115 [FREE Full text] [doi: 10.2196/11115] [Medline: 30135053]
7. Provisional Death Counts for Coronavirus Disease 2019 (COVID-19). Centers for Disease Control and Prevention. 2021 Oct 29. URL: https://www.cdc.gov/nchs/nvss/vsrr/covid19/index.htm [accessed 2021-10-30]
8. Sanders JM, Monogue ML, Jodlowski TZ, Cutrell JB. Pharmacologic treatments for Coronavirus Disease 2019 (COVID-19): A review. JAMA 2020 May 12;323(18):1824-1836. [doi: 10.1001/jama.2020.6019] [Medline: 32282022]
9. RECOVERY Collaborative Group. Dexamethasone in hospitalized patients with Covid-19. N Engl J Med 2021 Feb 25;384(8):693-704. [doi: 10.1056/nejmoa2021436]
10. Kaddoura M, AlIbrahim M, Hijazi G, Soudani N, Audi A, Alkalamouni H, et al. COVID-19 therapeutic options under investigation. Front Pharmacol 2020 Aug 6;11:1196 [FREE Full text] [doi: 10.3389/fphar.2020.01196] [Medline: 32848795]
11. Bai C, Chotirmall SH, Rello J, Alba GA, Ginns LC, Krishnan JA, et al. Updated guidance on the management of COVID-19: from an American Thoracic Society/European Respiratory Society coordinated International Task Force (29 July 2020). Eur Respir Rev 2020 Sep 30;29(157):200287 [FREE Full text] [doi: 10.1183/16000617.0287-2020] [Medline: 33020069]
12. Matthay MA, Thompson BT. Dexamethasone in hospitalised patients with COVID-19: addressing uncertainties. The Lancet Respiratory Medicine 2020 Dec;8(12):1170-1172. [doi: 10.1016/s2213-2600(20)30503-8]
13. LABEL: dexamethasone- Dexamethasone tablet dexamethasone- Dexamethasone solution, concentrate dexamethasone- Dexamethasone solution. U.S. National Library of Medicine. URL: https://dailymed.nlm.nih.gov/dailymed/drugInfo.cfm?setid=537b424a-3e07-4c81-978c-1ad99014032a [accessed 2021-10-30]
14. Boulware DR, Pullen MF, Bangdiwala AS, Pastick KA, Lofgren SM, Okafor EC, et al. A randomized trial of hydroxychloroquine as postexposure prophylaxis for Covid-19. N Engl J Med 2020 Aug 06;383(6):517-525 [FREE Full text] [doi: 10.1056/NEJMoa2016638] [Medline: 32492293]
15. Shah S, Das S, Jain A, Misra DP, Negi VS. A systematic review of the prophylactic role of chloroquine and hydroxychloroquine in coronavirus disease - 19 (COVID - 19). Int J Rheum Dis 2020 Apr 27;23(5):613-619. [doi: 10.1111/1756-185x.13842]
16. Liu M, Caputi TL, Dredze M, Kesselheim AS, Ayers JW. Internet searches for unproven COVID-19 therapies in the United States. JAMA Intern Med 2020 Aug 01;180(8):1116-1118 [FREE Full text] [doi: 10.1001/jamainternmed.2020.1764] [Medline: 32347895]
17. Petersen E, Koopmans M, Go U, Hamer DH, Petrosillo N, Castelli F, et al. Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics. The Lancet Infectious Diseases 2020 Sep;20(9):e238-e244. [doi: 10.1016/s1473-3099(20)30484-9]
18. WHO discontinues hydroxychloroquine and lopinavir/ritonavir treatment arms for COVID-19. World Health Organization. 2020 Jul 04. URL: https://www.who.int/news/item/04-07-2020-who-discontinues-hydroxychloroquine-and-lopinavir-ritonavir-treatment-arms-for-covid-19 [accessed 2021-10-30]
19. Horby PW, Mafham M, Bell JL, Linsell L, Staplin N, Emberson J, et al. Lopinavir–ritonavir in patients admitted to hospital with COVID-19 (RECOVERY): a randomised, controlled, open-label, platform trial. The Lancet 2020 Oct;396(10259):1345-1352. [doi: 10.1016/s0140-6736(20)32013-4]
20. Identify Geographical Location and Proxy by IP Address. IP2Location. URL: https://www.ip2location.com/ [accessed 2021-10-30]
21. SimilarWeb. URL: https://www.similarweb.com/ [accessed 2021-10-30]

22. Yang L, Lyons JG, Erickson SR, Wu C. Trends and characteristics of the US adult population's behavioral patterns in web-based prescription filling: National survey study. J Med Internet Res 2021 Mar 16;23(3):e23662 [FREE Full text] [doi: 10.2196/23662] [Medline: 33724193]
23. LABEL: Hydroxychloroquine sulfate tablet. U.S. National Library of Medicine. URL: https://dailymed.nlm.nih.gov/dailymed/drugInfo.cfm?setid=84b00366-96ef-41e1-bac6-5d24acdb9e1d [accessed 2021-10-30]
24. Ochsendorf F, Runne U. [Chloroquine and hydroxychloroquine: side effect profile of important therapeutic drugs]. Hautarzt 1991 Mar;42(3):140-146. [Medline: 2055762]
25. LABEL: KALETRA- lopinavir and ritonavir tablet, film coated KALETRA- lopinavir and ritonavir solution. U.S. National Library of Medicine. URL: https://dailymed.nlm.nih.gov/dailymed/drugInfo.cfm?setid=8290add3-4449-4e58-6c97-8fe1eec972e3 [accessed 2021-10-30]
26. Dryden L, Dewhurst N. Integration of a clinical pharmacist into a Canadian, urban emergency department: a prospective observational study. Int J Pharm Pract 2019 Apr;27(2):175-179. [doi: 10.1111/ijpp.12481] [Medline: 30079985]
27. Cicero TJ, Ellis MS. Health outcomes in patients using no-prescription online pharmacies to purchase prescription drugs. J Med Internet Res 2012 Dec 06;14(6):e174 [FREE Full text] [doi: 10.2196/jmir.2236] [Medline: 23220405]
28. Spector SL, Youdelman M. Analysis Of State Pharmacy Laws: Impact Of Pharmacy Laws On The Provision Of Language Services. National Health Law Program. 2010 Feb. URL: https://cdn.ymaws.com/www.wsparx.org/resource/resmgr/imported/PharmacyLawBookletFINAL.pdf [accessed 2021-10-30]
29. Mehta HB, Ehrhardt S, Moore TJ, Segal JB, Alexander GC. Characteristics of registered clinical trials assessing treatments for COVID-19: a cross-sectional analysis. BMJ Open 2020 Jun 09;10(6):e039978 [FREE Full text] [doi: 10.1136/bmjopen-2020-039978] [Medline: 32518212]
30. Gazendam A, Nucci N, Ekhtiari S, Gohal C, Zhu M, Payne A, et al. Trials and tribulations: so many potential treatments, so few answers. Int Orthop 2020 Aug 24;44(8):1467-1471 [FREE Full text] [doi: 10.1007/s00264-020-04625-7] [Medline: 32447429]
31. Fact Sheet for Health Care Providers: Emergency Use Authorization (EUA) of Hydroxychloroquine Sulfate Supplied from the Strategic National Stockpile for Treatment of COVID-19 in Certain Hospitalized Patients. Food and Drug Administration. 2020 Apr 27. URL: https://www.fda.gov/media/136537/download [accessed 2021-10-30]
32. RECOVERY Collaborative Group. Effect of Hydroxychloroquine in Hospitalized Patients with Covid-19. N Engl J Med 2020 Nov 19;383(21):2030-2040. [doi: 10.1056/nejmoa2022926]
33. Fajgenbaum DC, Khor JS, Gorzewski A, Tamakloe M, Powers V, Kakkis JJ, et al. Treatments administered to the first 9152 reported cases of COVID-19: a systematic review. Infect Dis Ther 2020 Sep 27;9(3):435-449 [FREE Full text] [doi: 10.1007/s40121-020-00303-8] [Medline: 32462545]
34. Memmel LM, Miller L, Gardner J. Over-the-internet availability of hormonal contraceptives regardless of risk factors. Contraception 2006 Apr;73(4):372-375. [doi: 10.1016/j.contraception.2005.10.001] [Medline: 16531170]
35. Penley B, Chen H, Eckel SF, Ozawa S. Characteristics of online pharmacies selling Adderall. J Am Pharm Assoc (2003) 2021 Jan;61(1):e103-e109 [FREE Full text] [doi: 10.1016/j.japh.2020.07.022] [Medline: 32912756]
36. Kim SH, Ryu YJ, Cho N, Kim AE, Chang J. Prescription drug price paradox: cost analysis of Canadian online pharmacies versus US Medicare beneficiaries for the top 100 drugs. Clin Drug Investig 2017 Oct 22;37(10):957-963. [doi: 10.1007/s40261-017-0556-6] [Medline: 28735450]
37. Penley B, Minshew L, Chen HH, Eckel S, Ozawa S. Accessibility of low-cost insulin from illegitimate internet pharmacies: Cross-sectional study. J Med Internet Res 2022;24(1):e25855. [doi: 10.2196/25855]
38. FDA warns repackers distributing pharmaceutical ingredients, including opioids, for putting consumers at risk with significant violations of manufacturing quality standards. Food and Drug Administration. 2019 Jul 02. URL: https://www.fda.gov/news-events/press-announcements/fda-warns-repackers-distributing-pharmaceutical-ingredients-including-opioids-putting-consumers-risk [accessed 2021-10-30]
39. Ozawa S, Evans DR, Bessias S, Haynie DG, Yemeke TT, Laing SK, et al. Prevalence and estimated economic burden of substandard and falsified medicines in low- and middle-income countries: a systematic review and meta-analysis. JAMA Netw Open 2018 Aug 03;1(4):e181662 [FREE Full text] [doi: 10.1001/jamanetworkopen.2018.1662] [Medline: 30646106]
40. A study on the public health and socioeconomic impact of substandard and falsified medical products. World Health Organization. 2020 Apr 06. URL: https://www.who.int/publications/i/item/9789241513432 [accessed 2021-10-30]
41. Newton PN, Bond KC, Adeyeye M, Antignac M, Ashenef A, Awab GR, et al. COVID-19 and risks to the supply and quality of tests, drugs, and vaccines. The Lancet Global Health 2020 Jun;8(6):e754-e755. [doi: 10.1016/S2214-109X(20)30136-4]
42. Michael White C. Counterfeit drugs: A major issue for vulnerable citizens throughout the world and in the United States. J Am Pharm Assoc (2003) 2021 Jan;61(1):e93-e98. [doi: 10.1016/j.japh.2020.04.020] [Medline: 32471767]

Abbreviations

CIPARx: Canadian International Pharmacy Association
EUA: Emergency Use Authorization
FDA: Food and Drug Administration

Edited by G Eysenbach; submitted 03.02.21; peer-reviewed by J Appalasamy, YM Huang; comments to author 16.04.21; revised version received 09.08.21; accepted 18.09.21; published 16.02.22.

Please cite as:
Ozawa S, Billings J, Sun Y, Yu S, Penley B
COVID-19 Treatments Sold Online Without Prescription Requirements in the United States: Cross-sectional Study Evaluating Availability, Safety and Marketing of Medications
J Med Internet Res 2022;24(2):e27704
URL: https://www.jmir.org/2022/2/e27704
doi:10.2196/27704
PMID:34662286

Original Paper

Age- and Sex-Specific Differences in Multimorbidity Patterns and Temporal Trends on Assessing Hospital Discharge Records in Southwest China: Network-Based Study

Liya Wang[1*], MA; Hang Qiu[1,2*], PhD; Li Luo[3], PhD; Li Zhou[4], MPA

[1]Big Data Research Center, University of Electronic Science and Technology of China, Chengdu, China

[2]School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China

[3]Business School, Sichuan University, Chengdu, China

[4]Health Information Center of Sichuan Province, Chengdu, China

[*]these authors contributed equally

Corresponding Author:
Hang Qiu, PhD
School of Computer Science and Engineering
University of Electronic Science and Technology of China
No.2006, Xiyuan Ave, West Hi-Tech Zone
Chengdu, 611731
China
Phone: 86 28 61830278
Fax: 86 28 61830278
Email: qiuhang@uestc.edu.cn

Abstract

Background: Multimorbidity represents a global health challenge, which requires a more global understanding of multimorbidity patterns and trends. However, the majority of studies completed to date have often relied on self-reported conditions, and a simultaneous assessment of the entire spectrum of chronic disease co-occurrence, especially in developing regions, has not yet been performed.

Objective: We attempted to provide a multidimensional approach to understand the full spectrum of chronic disease co-occurrence among general inpatients in southwest China, in order to investigate multimorbidity patterns and temporal trends, and assess their age and sex differences.

Methods: We conducted a retrospective cohort analysis based on 8.8 million hospital discharge records of about 5.0 million individuals of all ages from 2015 to 2019 in a megacity in southwest China. We examined all chronic diagnoses using the ICD-10 (International Classification of Diseases, 10th revision) codes at 3 digits and focused on chronic diseases with ≥1% prevalence for each of the age and sex strata, which resulted in a total of 149 and 145 chronic diseases in males and females, respectively. We constructed multimorbidity networks in the general population based on sex and age, and used the cosine index to measure the co-occurrence of chronic diseases. Then, we divided the networks into communities and assessed their temporal trends.

Results: The results showed complex interactions among chronic diseases, with more intensive connections among males and inpatients ≥40 years old. A total of 9 chronic diseases were simultaneously classified as central diseases, hubs, and bursts in the multimorbidity networks. Among them, 5 diseases were common to both males and females, including hypertension, chronic ischemic heart disease, cerebral infarction, other cerebrovascular diseases, and atherosclerosis. The earliest leaps (degree leaps ≥6) appeared at a disorder of glycoprotein metabolism that happened at 25-29 years in males, about 15 years earlier than in females. The number of chronic diseases in the community increased over time, but the new entrants did not replace the root of the community.

Conclusions: Our multimorbidity network analysis identified specific differences in the co-occurrence of chronic diagnoses by sex and age, which could help in the design of clinical interventions for inpatient multimorbidity.

(J Med Internet Res 2022;24(2):e27146) doi:10.2196/27146

KEYWORDS

multimorbidity pattern; temporal trend; network analysis; multimorbidity prevalence; administrative data; longitudinal study; regional research

Introduction

With the recent improvements in clinical interventions, advances in public health, lifestyle changes, and environmental exposures, multimorbidity has been a growing global health challenge [1-3]. Although multimorbidity is widely considered as the norm, not the exception, it still has an inconsistent definition and heterogeneity in methodology, which makes it difficult to gauge its prevalence and pattern in the general population [4-6]. In light of the increased mortality, lower quality of life, and higher utilization of health care services associated with multimorbidity [7-11], a global understanding of the multimorbidity pattern and trend is needed. Although a variety of studies have investigated the patterns of multimorbidity [12-16], most of them were conducted using cross-sectional surveys, which were generally limited either by their small number of self-reported conditions or by a small sample size. Therefore, a multidimensional approach is still needed to understand the full spectrum of multimorbidity networks, time trends, and patterns in age and sex, particularly in developing countries or regions [17].

With the enhancement of the storage capacity and accessibility of electronic information systems, digitized clinical record keeping has made routinely collected administrative data of unprecedented depth and variability available to researchers. This provides an opportunity for the application of network analysis to extract conceptual insights from large and messy data sets [18-20]. Although notable studies are few and mostly carried out in developed countries, they have provided promising findings in human phenotypic multimorbidity networks. For instance, based on the disease history of more than 30 million patients collected from hospital claims, correlations for more than 10,000 comorbid disease pairs were calculated and visualized in a phenotypic disease network [19]. Differences in identified multimorbidity across sex and racial groups were identified through macro analyses at the organ level [21,22]. Additionally, a study in Taiwan constructed an epidemiological disease network and examined its temporal pattern [23]. However, it is difficult to investigate the true extent of multimorbidity associations from these studies because of the differential definition of multimorbidity at the cross-sectional level or over a lifetime period, the difference in the measurement of associations, and the study settings that were mainly dominated by developed countries or regions.

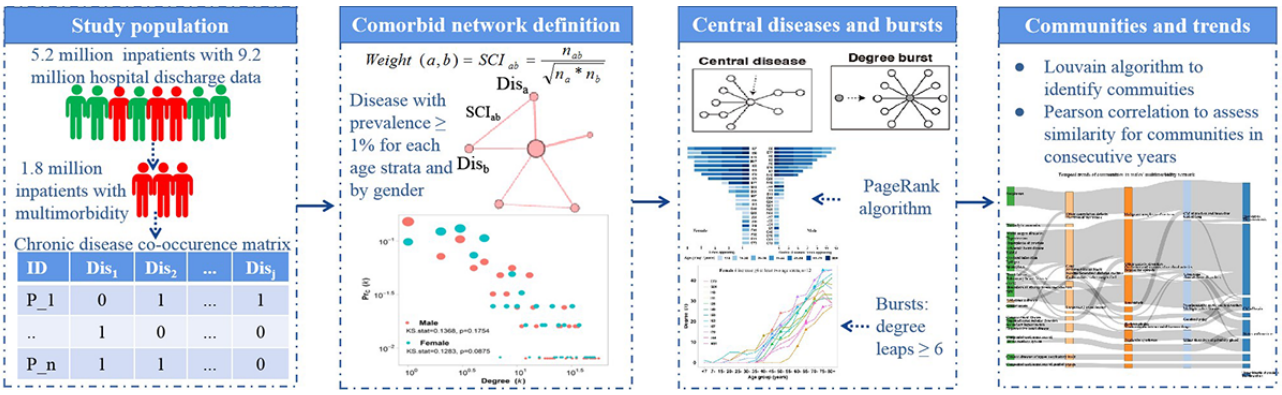
To address these gaps, we performed a retrospective study based on all inpatients living in a megacity in southwest China. We applied the standardized definition and classification system of multimorbidity [24]. Our major aim is to provide a multidimensional approach to understand the complex comorbid relationships among the full spectrum of chronic diseases in general multimorbidity inpatients in southwest China. Furthermore, this study aimed to assess age and sex differences in the multimorbidity pattern and investigate highly correlated communities and their temporal trends.

Methods

Overview

The workflow of this study is shown in Figure 1. First, we assessed the quality of the data set and confirmed the study population. Then, the cosine index was selected to construct sex- and age-specific multimorbidity networks. Next, we identified central diseases and bursts, and examined their differences across sex and age. Finally, we divided the networks into communities and assessed their temporal trends. Below, we provide more details on each step of the analysis.

Figure 1. Study workflow. SCI: Salton cosine index. Dis: chronic disease.



Ethics Approval

This study was approved by the Ethics Committee of the Health Information Center of Sichuan Province. The data were analyzed anonymously to maintain the privacy of inpatient data. As a study of previously collected administrative data, this work was exempt from informed consent requirements.

Data Source and Study Design

In this retrospective cohort analysis, we used the regional database of longitudinal clinical data for inpatients, which was provided by the Health Information Center of Sichuan Province. This regional database includes the anonymized hospital discharge reports (HDRs) collected from all of the 534

secondary hospitals and 144 tertiary hospitals in Sichuan Province; therefore, each inpatient's longitudinal clinical data were available. Each HDR contained information on the anonymized identity, age, sex, residential address, visit and discharge dates, principal discharge diagnosis, and up to 15 secondary diagnoses. All diseases were specified according to the ICD-10 (International Classification of Diseases, 10th revision) codes at 3 digits.

The eligibility criteria included inpatients who were residents of Chengdu and alive for the entire study period. A total of 5.2 million individuals (about 31.5% of Chengdu's population) with 9.2 million HDRs from 2015 to 2019 were included. As we were interested in diseases (ICD-10: A00-R99), hospitalizations in which the patients were marked only for general symptoms [24] (226,193 cases in total) were removed. According to the sex-specific diagnoses [21,24], 2329 male inpatients and 31 female inpatients were further removed due to conflicts between diagnoses and sex. Finally, the data preprocessing resulted in a total of 8.8 million hospitalizations corresponding to about 5.0 million individuals of all ages, and the sample was large enough to estimate age- and sex-specific multimorbidity patterns.

Network-Based Analysis

Chronic Diseases and Multimorbidity Definition

In 2018, the Academy of Medical Sciences recommended the adoption of a uniform definition and reporting system for multimorbidity [25], which identified multimorbidity as the co-existence of two or more chronic conditions (a physical noncommunicable disease, a mental health condition, or an infectious disease of long duration). Since chronic conditions would not be expected to go away in a single hospitalization period, we considered a 5-year period [13] rather than a single hospitalization for the definition of multimorbidity. The Chronic Condition Indicator [26], developed as part of the Healthcare Cost and Utilization Project, was used to differentiate between acute and chronic ICD-10 codes at 3 digits [24]. A total of 489 and 505 chronic disease codes were separately retained in males and females, respectively.

In order to generate more consistent and reliable estimates, we focused on chronic diseases with ≥1% prevalence for each of the following age strata: <7, 7-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, and 80+ years, and for both males and females [20], resulting in a total of 149 and 145 chronic diseases, respectively, which were further used in downstream analyses (Multimedia Appendix 1).

Multimorbidity Network Generation and Network Properties Calculation

A multimorbidity network developed from inpatients contains a set of nodes that are connected through edges. The node represents a chronic disease (ICD-10 codes at 3 digits), such that the node size is proportional to the disease prevalence and its color identifies the ICD-10 category.

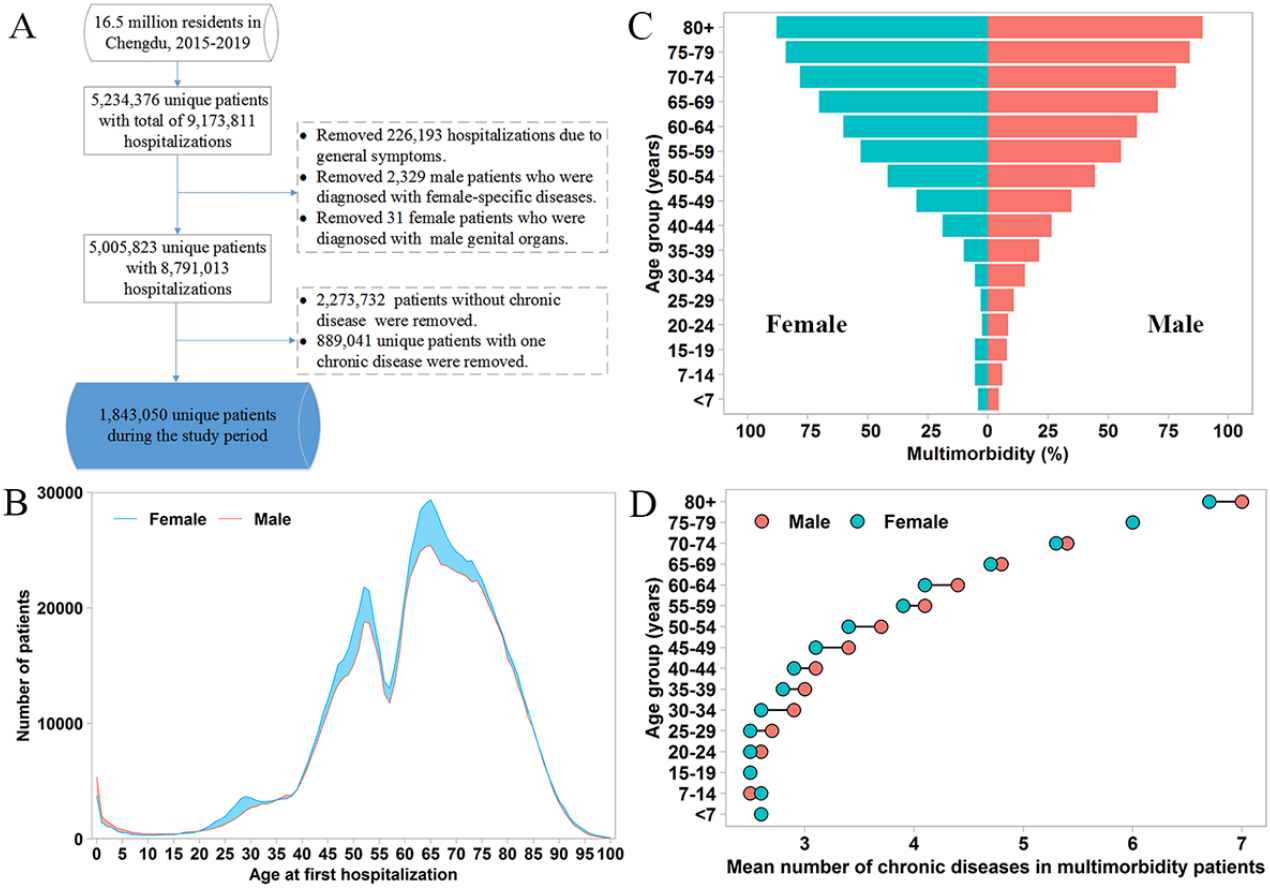
The edge in the multimorbidity network denotes the comorbid strength between co-existence diseases. Typically, the higher the comorbid strength of a disease pair, the lower the probability of co-existence by chance alone [19,20,27]. The relative risk (RR; calculated in Equation 1) or Pearson correlation coefficient (φ, calculated in Equation 2) was often used to quantify the comorbid strengths of disease pairs [19,20,27]. These 2 measures are not entirely independent of each other, as they are both affected by the sample size and have intrinsic bias [19]. As we were interested in tightly interconnected disease pairs, mutually exclusive disease pairs with negative comorbid strengths (RR < 1 or φ < 0) were excluded. Since the Salton cosine index (SCI; calculated in Equation 4) is immune to the sample size and only considers the co-occurrences and the prevalence of multimorbidity [28], we selected it to construct and compare the multimorbidity networks with a widely varied sample size in each of the sex- and age-specific groups.

where n_{ab} denotes the number of co-occurrences of diseases a and b, n_a and n_b represent the number of occurrences of diseases a and b, respectively, and N_{total} is the total number of inpatients in the sex/age-specific group.

Generally, a cutoff for the SCI is defined by assessing the relationship between the Pearson correlation and SCI, where the number of significantly correlated diseases is equal in both networks [29]. For each sex- and age-specific stratum, the steps to find a cutoff for the SCI are as follows: Step 1, Calculate the Pearson correlation coefficient (φ, calculated in Equation 2) and select the statistically significant correlations at α=.01 (calculated in Equation 3); Step 2, Find the minimum number of disease pairs  , where p, the maximum number of edges possible among n nodes detected in step 1, is equal to $n(n\text{-}1)/2$; Step 3, Find the number of pairs (q), where $n_{ab} \geq n_{ab\_minimum}$; Step 4, Find the SCI cutoff (Figure 2B), where the number of pairs is equal to q, detected in step 3. The above steps were used to create networks for males and females in each of the 16 age groups and then merge the same edges in different age groups into the general networks for males and females.

The Kolmogorov-Smirnov test was applied to investigate whether the degree distribution follows a power law. The structural properties can be measured using several network indices, such as the density, diameter, average path length, degree, weighted degree, closeness centrality, and betweenness centrality [30]. The closeness centrality measures the shortest distance of the disease away from other chronic diseases. Hence, the higher the closeness centrality of a disease, the higher the risk of co-occurrence with different diseases in fewer number of steps. The betweenness centrality denotes the number of shortest paths through a disease. Then, the higher the betweenness centrality of a disease, the higher the likelihood to form bridges between other diseases.

Figure 2. Characteristics of the study population. (A) Selection flow of the study population with multimorbidity. (B) Age and sex distribution for 1.8 million unique inpatients with multimorbidity. (C) Age- and sex-specific percentages of inpatients with multimorbidity among 5.0 million unique inpatients. (D) Age- and sex-specific mean numbers of chronic diseases among 1.8 million unique inpatients with multimorbidity.



The Central Diseases, Hubs, and Bursts

In order to distinguish the node centrality in the network, the PageRank algorithm [31] was applied, which considers the edge weights. The higher the PageRank value, the more "central" the disease [32]. The parameters were set as commonly assumed, where epsilon = 0.001 and probability = 0.85. Since no established guideline exists for how many nodes are central and since the number of nodes hugely differed among all the age groups in our study, we defined the central diseases as the nodes with the top 10 percentile of the PageRank value across the 16 age strata of males and females.

The connectivity of disease *a* is defined as the sum of all weights of all edges attached to it, which quantifies how strongly a disease is connected to the others. Diseases with larger connectivity are more likely to have a "system-wide" impact on the network. In the study, diseases with the top 10 percentile connectivity values for each of the 16 age strata of males and females have been referred to as the corresponding hubs.

In order to find the nodes with a vastly increased number of edges across age groups (defined as bursts) and to explore the sex difference in the age where the first large leaps occurred, we separately constructed male and female age-based trajectories of degree (k) for each node. The nodes with degree leaps ≥6 in the consecutive age groups with such leaps appearing at least one time in the subsequent stratum were defined as bursts. These bursts play essential roles in increasing the multimorbidity burden. Therefore, detecting the age of the first large leaps can help to understand the progression of multimorbidity.

Community Detection and Temporal Trends of Communities

Community detection separates the nodes of a generic undirected network into communities, such that connections within communities are stronger than those between them [33,34]. In order to identify distinct clusters of co-occurring diseases, we applied the Louvain method, a heuristic method based on modularity optimization [35]. Modularity Q is widely used to compare the partition quality and as an objective function to be optimized [35]. Furthermore, the community detection algorithm used in our analysis considers the weight of the links. Eigenvector centrality measures the influence of a node in a network [36]. The node with the largest eigenvector centrality in the community was therefore considered as the community root.

Observing how communities change over time can also provide valuable information about the network [23]. We applied the same methodology year by year and compared the results across time. As a result, we obtained the temporal trends of the multimorbidity networks. The Pearson correlation coefficient was used to measure the correlation for communities received in consecutive years.

All statistical analyses, network constructions, and visualizations were conducted in R software (version 3.5.1; R Development Core Team).

Results

Chronic Diseases and the Prevalence of Multimorbidity

About 5.0 million unique inpatients (representing about 30.3% of the overall Chengdu population) were enrolled in the study, among which 36.8% (a total of 1,843,050 unique inpatients, about 11.2% of the Chengdu population) had two or more chronic diseases (Figure 2). Demographically, the 1.8 million unique inpatients with multimorbidity consisted of inpatients of all ages with a higher percentage of females (52.1%). Generally, males had a statistically higher percentage of multimorbidity compared with females, except for the age group of 70-79 years (as shown in Multimedia Appendix 2). In addition, males in the middle-age (30-34 years and 45-64 years) and older elderly (80+ years) age groups had a large number of chronic diseases compared with females.

Properties of Age- and Sex-Specific Multimorbidity Networks

The phenotypic multimorbidity network analysis identified the network's global structure and uncovered chronic diseases with a closer co-occurrence (Figure 3). The RR and Pearson correlation coefficient used to measure disease co-occurrence are not entirely independent of each other (Figure 3A). Therefore, the SCI was used to measure the strength of comorbid diseases, and the cutoff of the SCI was determined by assessing the relationship between the Pearson correlation coefficient and SCI (Figure 3B). The cumulative distribution of the number of edges by nodes (degree (k) distribution) presented exponential decays (Figure 3C). Both the multimorbidity networks of males and females were scale-free as the distribution followed a power law (Kolmogorov-Smirnov test, P=.18 in the male network and P=.09 in the female network). The number of nodes and edges for the multimorbidity networks across age strata and by sex ranged from 22 to 74 and 18 to 579, respectively (Figure 3D and 3E). For patients above the age of 30 years, the number of edges we found was more significant in the male multimorbidity network. The number of edges became smaller in the lower age groups, but stronger disease connections were identified (Figure 3F). Table 1 lists the topological properties of each network. Generally, the multimorbidity networks in the younger age groups (≤40 years old) were sparser, except for females <7 years old. The maximum diameter and average path length of the female multimorbidity network were 10 and 4, respectively, which were larger than those of the male network (8 and 2.8, respectively). The average closeness centrality in the multimorbidity network of middle-aged (30-54 years) males was significantly higher than that for females (Wilcoxon test, both P<.05).

Figure 3. The properties of sex-specific phenotypic multimorbidity networks. (A) Scatter plot between the relative risk (RR) and the Pearson correlation coefficient of disease pairs; due to the interest of tightly interconnected diseases, we excluded mutually exclusive disease pairs with RR <1 or correlation <0. (B) The cutoff of the Sclton cosine index (SCI) where the numbers of significant disease pairs are equal in networks using the Pearson correlation coefficient and SCI. (C) Degree (k) distributions for sex-specific multimorbidity networks using the SCI. (D) and (E) The numbers of connected nodes and edges in each multimorbidity network across age strata and by sex. (F) Box plot of the SCI across age strata and by sex. The width of the box is proportion to the number of edges in each strata's network.

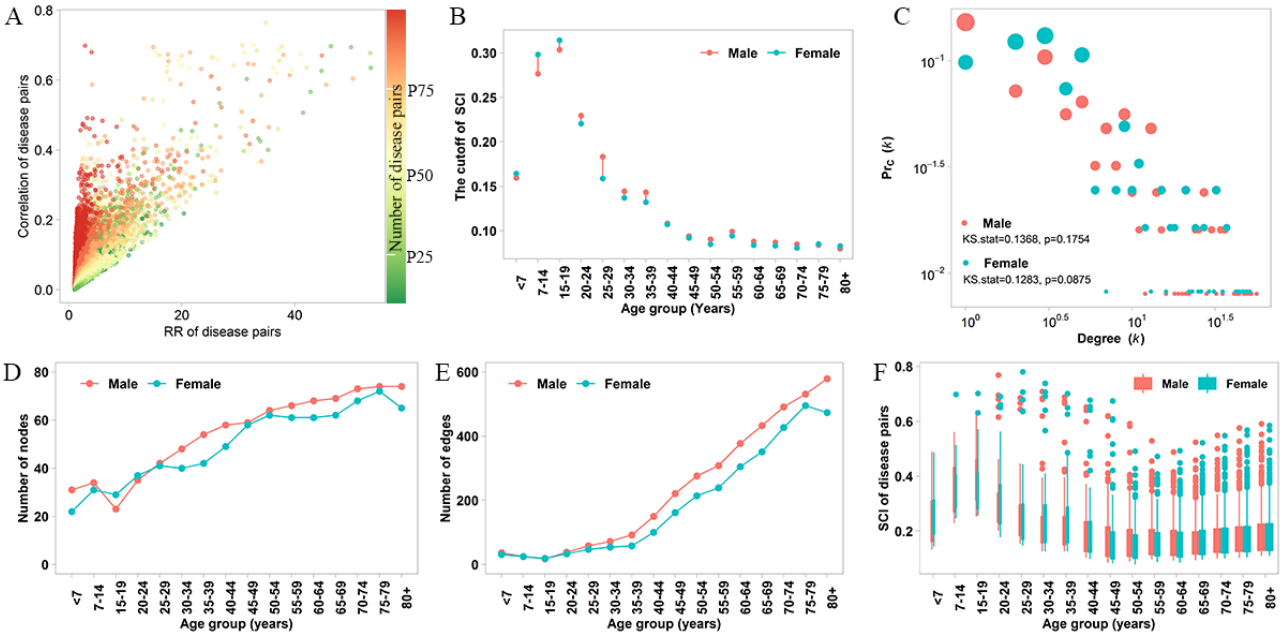


Table 1. Sex- and age-specific multimorbidity network properties.

Network in each age group (years)	Density	Diameter	Average path length	Average degree	Avg.w degree[a]	Avg.clos centrality[b]	Avg.bet centrality[c]
Male multimorbidity network							
<7	0.077	4	1.6	2.3	0.58	0.74	4.4
7-14	0.045	4	1.5	1.5	0.50	0.82	2.4
15-19	0.071	4	1.5	1.6	0.62	0.86	3.0
20-24	0.064	4	1.9	2.2	0.71	0.67	7.3
25-29	0.067	5	2.4	2.8	0.75	0.57	16.5
30-34	0.064	8	2.5	3.0	0.73	0.52[d]	23.4
35-39	0.064	6	2.4	3.4	0.78	0.51[d]	20.4
40-44	0.091	6	2.8	5.2	1.02	0.40[d]	57.8
45-49	0.129	6	2.7	7.5[d]	1.30[d]	0.38[d]	63.5
50-54	0.137	5	2.6	8.6[d]	1.47	0.39[d]	73.6
55-59	0.144	5	2.5	9.3	1.61	0.41	64.5
60-64	0.165	5	2.4	11.1	1.84	0.43	62.1
65-69	0.185	6	2.3	12.6	2.13	0.45	64.0
70-74	0.187	5	2.1	13.5	2.40	0.48	58.6
75-79	0.197	5	2.1	14.4	2.63	0.49	61.8
80+	0.214	4	1.9	15.6	3.00	0.53	50.2
Female multimorbidity network							
<7	0.134	2	1.3	2.8	0.74	0.83	4.0
7-14	0.054	3	1.4	1.6	0.58	0.82	2.3
15-19	0.047	2	1.2	1.3	0.50	0.91	1.7
20-24	0.050	4	1.8	1.8	0.59	0.72	4.6
25-29	0.057	5	2.1	2.3	0.64	0.61	8.7
30-34	0.069	8	3.4	2.7	0.70	0.42	45.5
35-39	0.067	10	4.0	2.8	0.70	0.36	69.2
40-44	0.085	9	3.7	4.1	0.82	0.34	88.7
45-49	0.098	6	3.1	5.6	0.93	0.34	94.3
50-54	0.113	7	2.9	6.9	1.08	0.36	82.2
55-59	0.131	6	2.7	7.8	1.30	0.39	68.7
60-64	0.167	5	2.4	10.0	1.65	0.43	59.4
65-69	0.186	5	2.3	11.3	1.96	0.45	50.2
70-74	0.187	5	2.2	12.6	2.28	0.48	58.5
75-79	0.194	5	2.1	13.8	2.57	0.50	57.2
80+	0.227	4	2.0	14.6	2.80	0.53	47.3

[a]Avg.w degree: average weighted degree.

[b]Avg.clos centrality: average closeness centrality.

[c]Avg.bet centrality: average betweenness centrality.

[d]The values in the multimorbidity network in males were statistically higher than those in females (P<.05).

Age- and Sex-Specific Differences in Central Diseases, Hubs, and Bursts

The female and male multimorbidity networks are visualized in Figure 4A and 4B, respectively. According to the frequency and comorbidity strength, the top 20 comorbid disease pairs involved 13 diseases, among which 11 diseases (E11, E78, I10, I11, I25, I27, I50, I63, I67, I70, and J44) were common to both males and females, and 2 diseases were sex specific (hyperplasia of the prostate [N40] in males and spondylosis [M47] in females). The most comorbid disease pair was essential hypertension (I10) with cerebral infarction (I63), which occurred in males older than 30 years and females older than 40 years. Notably, there existed a few disease pairs that exhibited strong comorbid strengths but only occurred in a typical age group, for instance, congenital malformation co-existence in children <7 years old. Based on the PageRank algorithm, 23 and 26 chronic diseases were identified as central diseases in the female and male multimorbidity networks. Among these diseases, 14 chronic diseases were common to both males and females, and comprised critical diseases across different ages, such as heart failure (I50), essential hypertension (I10), glycoprotein metabolism disorders (E77), and lipoprotein metabolism disorders (E78) (Figure 4C). Interestingly, depressive episodes (F32) and other anxiety disorders (F41) represented the central diseases among females aged 7-14 years and 25-29 years, respectively. A total of 26 unique diseases were hubs in the multimorbidity networks, including 19 hubs common to both males and females, 1 female-specific hub (spondylosis at 50-59 years), and 6 male-specific hubs (Figure 4D). Furthermore, for each burst, which had at least 2 degree leaps ≥6 through consecutive age groups, connectivity trajectories across age groups are presented in Figure 4E and 4F. A total of 7 burst nodes were common to both males and females, among which essential hypertension (I10) first occurred in men aged 30 to 34 years. Among the 4 male-specific burst nodes, the earliest leaps were glycoprotein metabolism disorders (E77), which happened at 25-29 years. Remarkably, 9 diseases were classified as not only central diseases, but also hubs and bursts. Among them, 5 were common to both males and females, including essential hypertension (I10), chronic ischemic heart disease (I25), cerebral infarction (I63), other cerebrovascular diseases (I67), and atherosclerosis (I70). Therefore, particular diseases act as both bursts for increasing the network complex and as hubs for having a “system-wide” impact on the network, and some of these act as central diseases for playing the most important role in the network.

Disease progression by age was evaluated by analyzing the connectivity trajectories of each disease (Figure 5). Males had a higher connectivity, except for the youngest age groups (≤14 years old). In contrast, females had a steeper slope, particularly those aged 55+ years. The central diseases in both males and females showed a higher connectivity compared with noncentral diseases, and the connectivity difference between central diseases and noncentral diseases increased with age among males and females older than 35 years. A similar pattern also appeared in the hubs, but its connectivity difference was more conspicuous at a younger age compared with that in central diseases, which is consistent with the network topology. Subsequently, the bursts in males and females had a higher connectivity after the initial degree leaps ≥6 occurred (males older than 35 years and females older than 40 years).

Figure 4. Multimorbidity networks, central diseases, hubs, and connectivity trajectories. Age-adjusted multimorbidity network in females (A) and males (B). Nodes represent chronic diseases (ICD-10 [International Classification of Diseases, 10th revision] codes at 3 digits), such that the node size is proportional to the disease prevalence among multimorbidity patients and its color identifies the ICD-10 category. Link weights are proportional to the magnitudes of the cosine index. (C) The central diseases in each age strata by sex. Diseases with the top 10 percentiles for PageRank in each strata were identified as central diseases. (D) The hubs in each age strata and by sex. Nodes with the top 10 percentiles for hubs in each strata were identified as hubs. The age-based trajectories of the degree (k) of bursts in females (E) and males (F). The triangles indicate degree leaps ≥6 through the consecutive age groups. The disease with at least two such degree leaps was defined as a burst, which means the bursts of disease associations leading to multimorbidity.

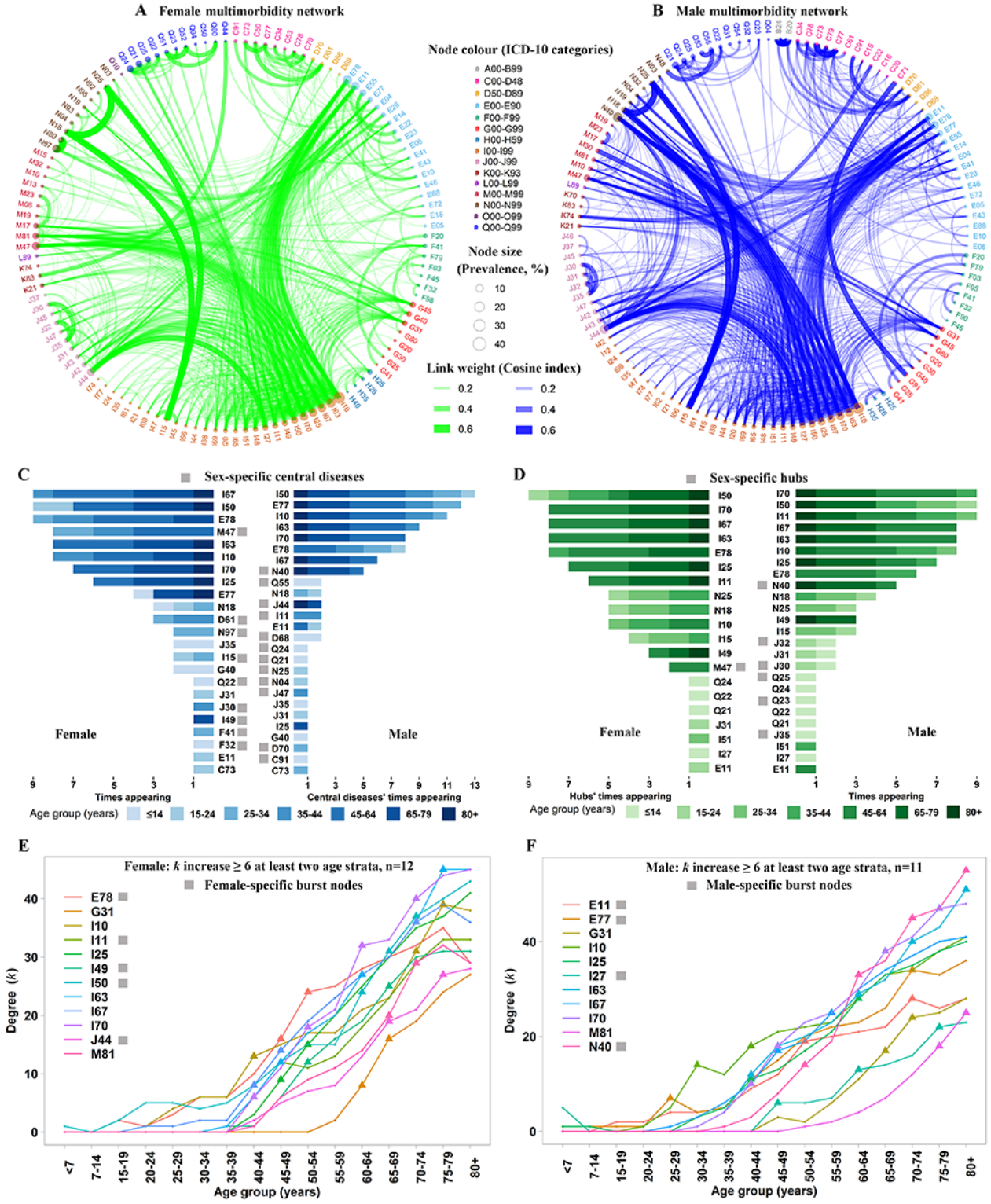
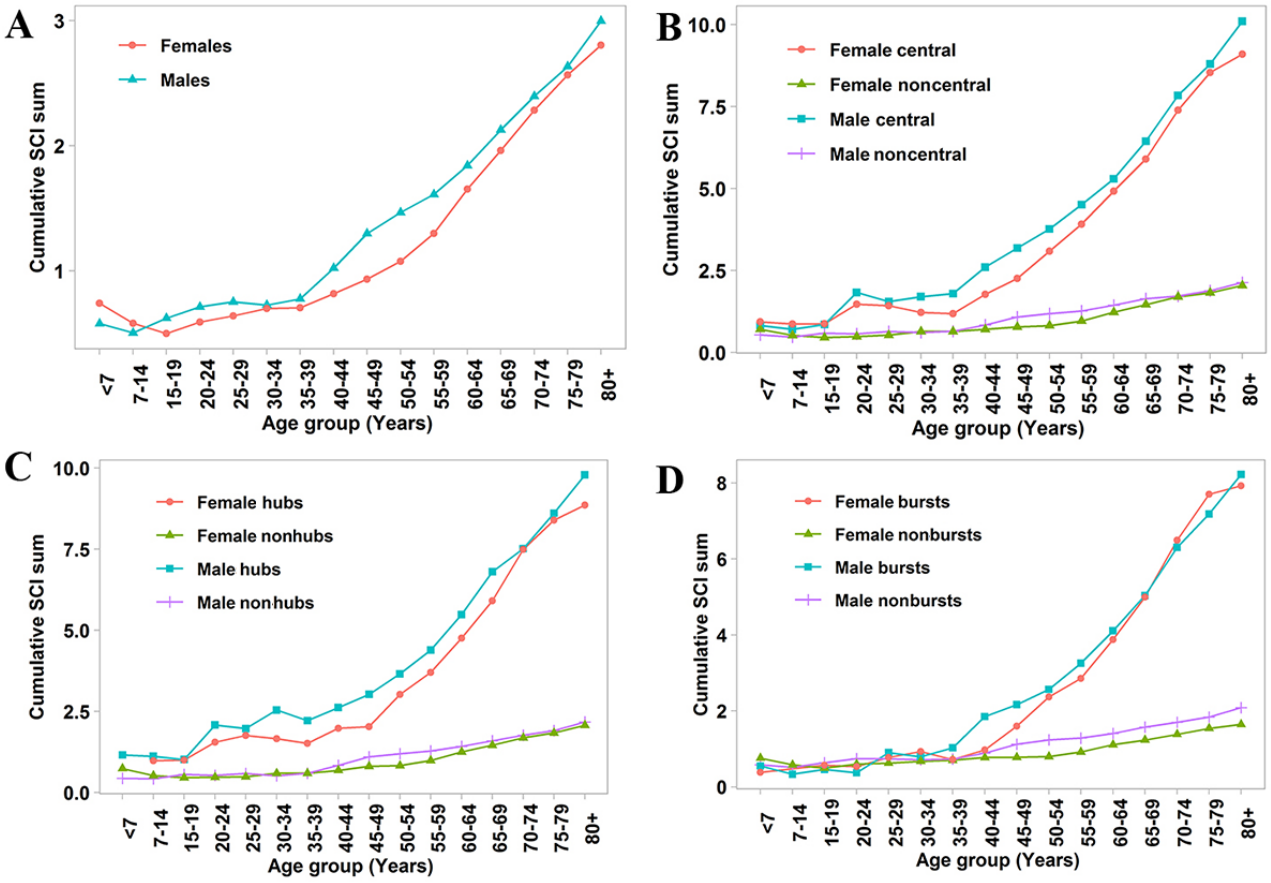


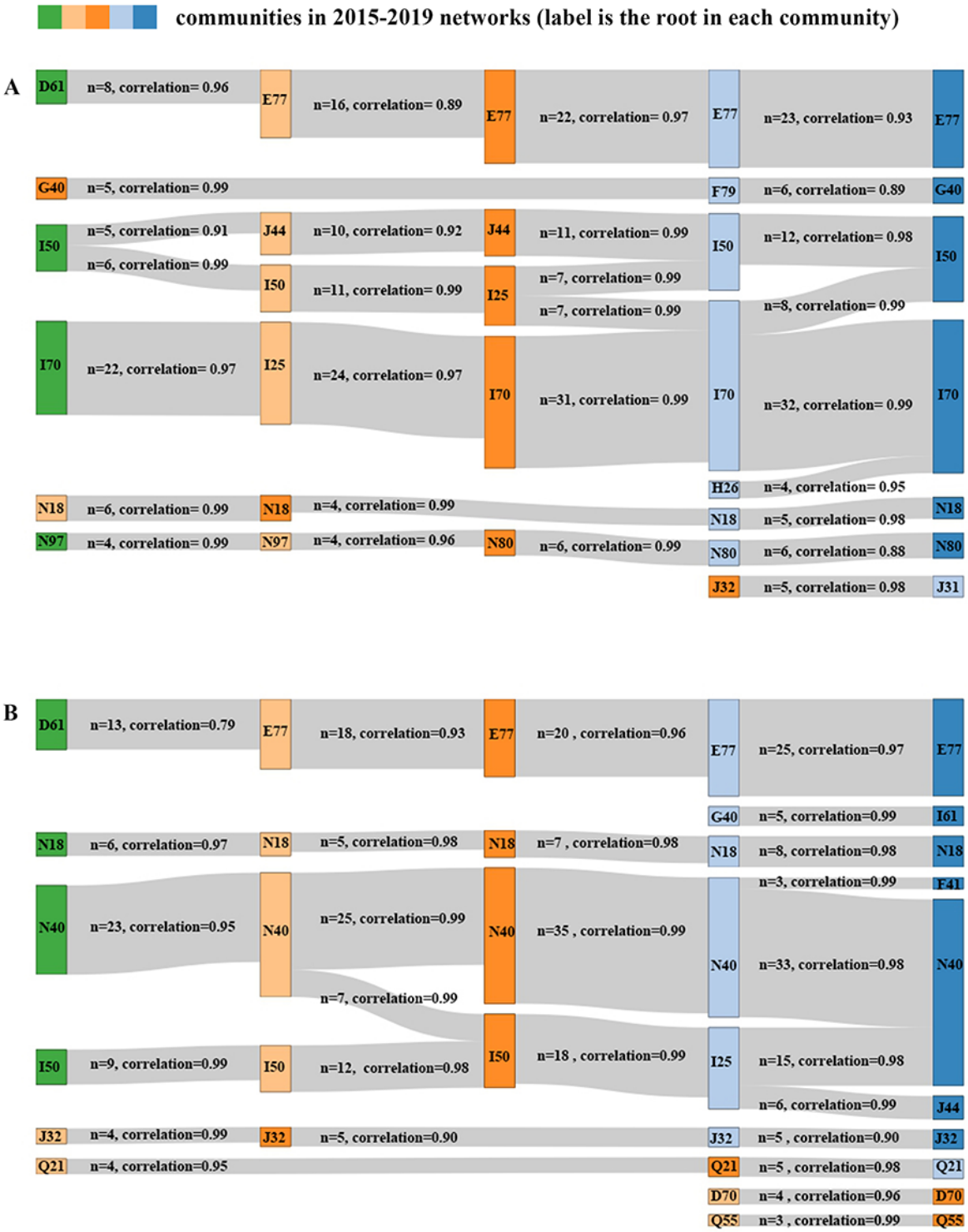
Figure 5. Sex-specific connectivity (cumulative of the node-average Sclton cosine index [SCI]) across age groups. All nodes (A), central diseases vs noncentral diseases (B), hubs vs nonhubs (C), and bursts vs nonbursts (D).



Temporal Trends of Communities

The community structures showed little variation across time, while the community root tended to be stable throughout time (Figure 6). The number of chronic diseases in the community increased over time, consistently in both males and females, and the new entrants did not replace the community root. For instance, the number of diseases in the community, among which the disorder of glycoprotein metabolism (E77) or other aplastic anemias (D61) was identified as the root, increased from 8 to 23 in the female community and from 13 to 25 in the male community. In addition, the community common to both males and females was defined as having the same root within the community, where many diseases were common to both sex groups and few diseases were sex specific. For instance, in the community with chronic renal failure (N18) as the root, both the female and male communities included the same diseases, such as secondary hypertension (I15), chronic nephritic syndrome (N03), nephrotic syndrome (N04), and disorders resulting from impaired renal tubular function (N25), while the male community also included vitamin D deficiency (E55) and the female community also included gout (M10) and systemic lupus erythematosus (M32). The clustering of mental health disorders, including depressive episode (F32), other anxiety disorders (F41), and somatoform disorders (F45), differed by sex. For example, the male community included only mental health disorders, while the female community included mental health disorders and various physical diseases. As for sex-specific diseases, the majority of female-specific diseases were in a separate community, with female infertility (N97) or endometriosis (N80) as the root. As for male-specific diseases, hyperplasia of the prostate (N40) was consistent throughout time, and its eigenvector centrality was even higher than that of atherosclerosis, heart failure, or cerebral infarction.

Figure 6. Temporal trends of communities in the female (A) and male (B) multimorbidity networks. By conducting networks year by year and comparing across time, we were able to obtain the temporal trends of the communities. The root, defined as the node with the highest eigenvector centrality within the community, is labeled using ICD-10 (International Classification of Diseases, 10th revision) at 3 digits. Similarity over time is assessed using the Pearson correlation coefficient for communities obtained in consecutive years, and unsignificant (*P*>.05) similarities are excluded. The "n" value is the number of chronic diseases consistent in the consecutive years.



Discussion

Principal Findings

We constructed multimorbidity networks among multimorbidity inpatients (about 11.2% of the Chengdu population) consisting of all ages, which established the connections between chronic diseases in the general hospitalized population from a megacity with 16.5 million residents in southwest China. Multimorbidity affected people of all ages, and their complex interactions were more intensive among males and inpatients ≥40 years old. Notably, in the female multimorbidity network, mental health

disorders co-occurred with various mental and physical diseases (eg, metabolic disorders, cardiovascular diseases, and neurodegeneration diseases), among which the co-existence of a depressive episode with other anxiety disorders was detected in 14 age groups of 7-79 years. Moreover, disease connectivity leaps, central diseases, and highly interlinked communities were detected. To the best of our knowledge, this is the first regional study within a developing country that applied regional hospital discharge records rather than self-reported survey data to provide an overview of the prevalence of multimorbidity, obtain multimorbidity patterns, and assess the sex and age differences. Our results demonstrate that the application of network-based algorithms to routinely collected health care data might provide a way to better screen and identify the complex interactions among chronic diseases.

Multimorbidity Affects People of All Ages

Multimorbidity affects people of all ages, even children who are inpatients (≤14 years old), where 5% of these children have at least two chronic diseases. The findings were in accordance with what was reported in previous studies based on the general population because the prevalence differed by age range, sex, ethnicity, socioeconomic status, and lifestyle, cultural, and health-seeking behaviors [6,13,37-39]. Our results showed that 36.8% of the inpatients living in Chengdu during the 2015-2019 time period had at least two chronic diseases, which is lower than that in the Netherlands (multimorbidity prevalence of 57%) [39], Spain (multimorbidity prevalence of 43.2%) [40], and Canada (multimorbidity prevalence of 53.3%) [6], but higher than that in the United Kingdom (multimorbidity prevalence of 19%) [14], Scotland (multimorbidity prevalence of 31.1%) [41], Singapore (multimorbidity prevalence of 26.2%) [42], Italy (multimorbidity prevalence of 15.3%) [43], and Denmark (multimorbidity prevalence of 21.6%) [44]. A scoping review found a wide range in multimorbidity prevalence in the general population as reported in studies using a large data set, from 15.3% to 68.4% [45]. The reported multimorbidity prevalence is still highly varied due to inconsistent measurements of chronic conditions and multimorbidity [45,46]. Additionally, a study using claims data in Beijing reported that the prevalence of multimorbidity was 51.6% and 81.3% for middle-aged adults (45-59 years) and older adults (≥60 years), respectively [47], which were higher than the respective prevalences of 41.7% and 75.2% in our study. One explanation might be related to the differential study design, since the study in Beijing used both outpatient and inpatient clinical diagnoses for the measurement of multimorbidity. However, the study in Beijing only used 13 most frequently mentioned diseases to measure multimorbidity and the study population was restricted to people who had been employed, which would limit its generalizability to a general population. Therefore, estimating the prevalence of multimorbidity based on a regional database is essential for the design of health care strategies. To the best of our knowledge, this is the first regional study within a developing country to provide an overview of the prevalence of multimorbidity using regional hospital discharge records rather than self-reported survey data. The prevalence of multimorbidity increased with age, which is in line with the findings of previous epidemiological studies that the prevalence of multimorbidity may be increasing, at least in part, because of population aging [5,6,8,41].

Age and Sex Differences in Multimorbidity Patterns

We identified the multimorbidity patterns in age- and sex-specific inpatient groups, which were comparable with previous studies in developed countries or regions [6,16,20,21]. For instance, Ioakeim-Skoufa et al [16] found associations of respiratory disorders with circulatory diseases, and depression and anxiety with chronic musculoskeletal diseases. In our study, we identified the most frequent and strongest comorbid disease pairs, such as the associations of circulatory diseases with endocrine diseases, diseases of the musculoskeletal system, and respiratory disorders, which co-occurred more frequently than expected by random chance. Our data set consisted of hospitalizations of all ages, and thus contained information about the diseases that are common and specific among age and sex groups. For example, we identified some disease pairs that occurred throughout life (eg, heart failure co-occurring with complications of heart disease and lipoprotein metabolism disorders co-occurring with diabetes mellitus) and some pairs with a stronger comorbid strength but only occurrence among a typical age group (eg, congenital malformation co-existence in children <7 years old). Intuitively, chronic diseases would be expected to co-occur in an individual if their resilience or vulnerability was altered or if they shared a common pattern of influence [48-50]. Thus, as in previous studies assessing the disease trajectory in patients with depression [51] and type 2 diabetes [52], and the general population [24], regional databases collecting HDRs spanning a sufficient time period (generally 10+ years as in the above mentioned studies) will support further studies to explore the potential causal directions among complex correlations. In accordance with previous studies [21,53], we identified a difference in the associations of mental disorders with physical diseases among different sex groups, and generally stronger associations were found in females than in males. This sex difference within mental health multimorbidity may be related to differences in the patient care–seeking behaviors between males and females, as indicated in a previous study that reported on how social factors can discourage males from seeking mental health care [54]. Our findings support the development of interdisciplinary and multidisciplinary treatment strategies for patients with depression or anxiety [53,55,56], since they frequently had physical diseases, such as metabolic disorders, Alzheimer disease, epilepsy, hypertension, chronic ischemic heart disease, heart failure, cerebral infraction, atherosclerosis, gastroesophageal reflux disease, decubitus ulcer, and spondylosis. The data-driven discovery of diseases co-occurring might be useful to generate potential hypotheses for coexisting diseases (eg, sharing the same gene, having common risk factors, and displaying a consistent temporal progression trend [57,58]) and their differences in age and sex (eg, physical, hormonal, and even genetic differences by sex [59,60] and disease progression with age [24]). Additionally, the data-driven discovery of diseases co-occurring, especially based on a complete population with a high-quality health care database, might have impacts on disease management [50].

We found lifelong comorbid disease pairs among multimorbidity inpatients, for example, lipoprotein metabolism disorders (E78)

co-occurring with diabetes mellitus (E11) in both males and females ≥15 years old, heart failure (I50) co-occurring with complications of heart disease (I51) in females ≥15 years old, and a depressive episode (F32) co-occurring with other anxiety disorders (F41) in females aged 7-79 years. Congenital malformations, generally the earliest diagnosed diseases in life (ie, pre- or perinatally), had a higher prevalence in multimorbid girls (<7 years) than in boys (<7 years) in our study, especially congenital malformations of the circulatory system (including congenital malformations of the cardiac septa, Q21; congenital malformations of the pulmonary and tricuspid valves, Q22; congenital malformations of the aortic and mitral valves, Q23; other congenital malformations of the heart, Q24; and congenital malformations of the great arteries, Q25). The prevalence of congenital heart disease in girls was higher than in boys [61], and treatment or progression of congenital heart disease could cause complications of heart disease and heart failure [62,63], which may support the finding in our study that heart failure co-occurring with complications of heart disease was earlier in females than in males. Chronic diseases with the earliest connectivity leaps begin at 25-29 years in males, about 15 years earlier than in females. These findings indicate the need to appropriately handle multimorbidity among youth or middle-aged patients [2,64]. Attention and activities are required to prevent such people from entering the multimorbidity category, especially for lipoprotein metabolism disorders, diabetes mellitus, hypertension, heart failure, spondylosis, chronic renal failure, fibrosis and cirrhosis of the liver, and gout. In addition, appropriate guidelines and flexible care management support systems are required across a broader age range.

Within each multimorbidity network, we identified the central disease, which played the most important role in the network (eg, having a large number of comorbid diseases and thus increasing the scale of network, having a relatively smaller number of comorbid diseases but exhibiting stronger comorbid strengths, and playing the role of connectivity to connect those unconnected diseases). In the biomedical scenario, central diseases may be interpreted as those that are more likely to appear in patients with multimorbidity or lead to multimorbidity. Therefore, common causal genes and molecular processes or signaling pathways may be shared among central diseases and their neighbors [27,65]. We found that circulatory diseases and metabolic diseases were the most important diseases among almost all age groups; thus, clinical studies of the identified central diseases may be helpful to improve prevention strategies and health care policies [2]. Notably, younger females should receive more attention on 2 mental health disorders, depressive episodes and other anxiety disorders, which significantly increased the scale of their multimorbidity network. The observations of central diseases could have important implications on the design of health care prevention, such that measurements targeting a specific factor may benefit many related diseases.

Furthermore, we observed that some communities remained stable across time, while others became more extensive due to the occurrence of more diseases. Few studies have observed the temporal trends of networks or communities [23,37]. Jiang et al found that the network structure, connectivity, and module structure varied across time [23]. The work of van Oostrom et al [66] showed that the prevalence of chronic diseases in the general practice registration over the period between 2004 and 2011 increased from 34.9% to 41.8%, and this increase could be only partially explained by the aging of the population. In our study, the community of mental health disorders in the male multimorbidity network consisted of depressive episodes, other anxiety disorders, and somatoform disorders, which seemed independent with physical diseases, while the community involving females additionally included various physical diseases, such as disorders of lipoprotein metabolism and other lipidemias (E78), essential hypertension (I10), hypertensive heart disease (I11), cerebral infarction (I63), atherosclerosis (I70), gastroesophageal reflux disease (K21), and spondylosis (M47). The clustering difference in mental health disorders according to sex might be related with the higher underdiagnosis rate of mental health disorders in males [67] and provide evidence for differential strategies in diagnosis and treatment for males and females. For instance, when 2 diseases are discordant in terms of their pathogenesis (eg, depression co-occurring with cerebrovascular disease in females), they may require separate time-intensive treatment plans [56,68-70]. Additionally, within the community, both concordant and discordant diseases in terms of their pathogenesis were included (eg, a female community with atherosclerosis as the root and including diabetes mellitus, hypertension, cerebrovascular diseases, mental health disorders, and spondylosis), which might lead to very different management needs and treatment strategies [56,68-71]. Communities can describe the interconnections among chronic diseases, with more tight connections between those in the same community. In a further study, it will be of interest to examine the direction of these interconnections or to explore their common risk factors for priority management.

Strengths and Limitations

The main strengths of this study can be summarized as follows. First, this is the first regional study in developing countries based on a large-scale data set (8.8 million hospital discharge records) to examine multimorbidity patterns and trends, and their differences across age and sex. Additionally, a network-based approach is applied to extract conceptual insights from routinely collected hospital discharge records. The use of this method can be extended to other health care data sets. Lastly, the use of routinely collected administrative data at a regional level is advantageous because the data are uniformity distributed and unbiased, which provides an opportunity to identify the co-occurrence of rare clinical diseases.

This study has some limitations. First, the main limitation of this study is the unavailability of individual-level socioeconomic status, lifestyle, and clinical variables. These factors would play essential roles in understanding the differences among multimorbidity patterns [23]. This limitation is common among studies that use routinely collected health care data sets. Second, the data set that was used did not contain information on the outpatients who were seeking solely outpatient care. Hence, it is vital to interpret our findings in the context of the inpatient population in a developing country. Third, we excluded individuals who were not alive during the study period to obtain a more homogenous study population, which may underestimate

the diseases that have high mortality rates. However, it was demonstrated by a previous study that this exclusion criterion did not drastically impact the results [37].

Conclusions

In this paper, we performed a network-based analysis of 8.8 million hospital discharge records and identified age and sex differences in multimorbidity patterns and the evolution of multimorbidity over time. This longitudinal study provides the first evidence from a developing country that multimorbidity affects people of all ages and their complex interactions are more intensive among males and inpatients ≥40 years old. Mental health disorders were comorbid with more various mental and physical diseases in females than in males. The lifelong comorbid disease pairs, disease connectivity leaps, central diseases, highly interlinked communities, and age- and sex-specific comorbidity patterns detected in the study might provide suggestions for enhancing integrated management in multimorbidity patients. Meanwhile, the network-based approach applied in our study could investigate all the multimorbidity connections at the population level, which could be used within health care data sets in other settings.

Acknowledgments

This work was supported by the Key Research and Development Project of Sichuan Province, China (grants 2018SZ0114 and 2019YFS0271).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Diseases with prevalence ≥1% in males and females.
[DOCX File , 33 KB - jmir_v24i2e27146_app1.docx]

Multimedia Appendix 2
Multimorbidity prevalence in each age strata and by sex.
[DOCX File , 17 KB - jmir_v24i2e27146_app2.docx]

References

1. Forman DE, Maurer MS, Boyd C, Brindis R, Salive ME, Horne FM, et al. Multimorbidity in Older Adults With Cardiovascular Disease. J Am Coll Cardiol 2018 May 15;71(19):2149-2161 [FREE Full text] [doi: 10.1016/j.jacc.2018.03.022] [Medline: 29747836]
2. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. The Lancet 2012 Jul 7;380(9836):37-43 [FREE Full text] [doi: 10.1016/S0140-6736(12)60240-2] [Medline: 22579043]
3. Calderón-Larrañaga A, Vetrano DL, Ferrucci L, Mercer SW, Marengoni A, Onder G, et al. Multimorbidity and functional impairment-bidirectional interplay, synergistic effects and common pathways. J Intern Med 2019 Mar;285(3):255-271 [FREE Full text] [doi: 10.1111/joim.12843] [Medline: 30357990]
4. Xu X, Mishra GD, Jones M. Evidence on multimorbidity from definition to intervention: An overview of systematic reviews. Ageing Res Rev 2017 Aug;37:53-68. [doi: 10.1016/j.arr.2017.05.003] [Medline: 28511964]
5. Griffith LE, Gilsing A, Mangin D, Patterson C, van den Heuvel E, Sohel N, et al. Multimorbidity Frameworks Impact Prevalence and Relationships with Patient-Important Outcomes. J Am Geriatr Soc 2019 Aug;67(8):1632-1640. [doi: 10.1111/jgs.15921] [Medline: 30957230]
6. Nicholson K, Terry AL, Fortin M, Williamson T, Bauer M, Thind A. Prevalence, characteristics, and patterns of patients with multimorbidity in primary care: a retrospective cohort analysis in Canada. Br J Gen Pract 2019 Sep;69(686):e647-e656 [FREE Full text] [doi: 10.3399/bjgp19X704657] [Medline: 31308002]
7. Multimorbidity: Understanding the challenge. Richmond Group of Charities. 2018. URL: https://richmondgroupofcharities.org.uk/sites/default/files/multimorbidity_-_understanding_the_challenge.pdf [accessed 2020-10-12]
8. Cassell A, Edwards D, Harshfield A, Rhodes K, Brimicombe J, Payne R, et al. The epidemiology of multimorbidity in primary care: a retrospective cohort study. Br J Gen Pract 2018 Apr;68(669):e245-e251 [FREE Full text] [doi: 10.3399/bjgp18X695465] [Medline: 29530918]
9. Payne RA, Abel GA, Guthrie B, Mercer SW. The effect of physical multimorbidity, mental health conditions and socioeconomic deprivation on unplanned admissions to hospital: a retrospective cohort study. CMAJ 2013 Mar 19;185(5):E221-E228 [FREE Full text] [doi: 10.1503/cmaj.121349] [Medline: 23422444]
10. Payne RA, Mendonca SC, Elliott MN, Saunders CL, Edwards DA, Marshall M, et al. Development and validation of the Cambridge Multimorbidity Score. CMAJ 2020 Feb 03;192(5):E107-E114 [FREE Full text] [doi: 10.1503/cmaj.190757] [Medline: 32015079]

11. Randall DA, Lujic S, Havard A, Eades SJ, Jorm L. Multimorbidity among Aboriginal people in New South Wales contributes significantly to their higher mortality. Med J Aust 2018 Jun 02;209(1):19-23. [doi: 10.5694/mja17.00878] [Medline: 29954311]
12. Khorrami Z, Rezapour M, Etemad K, Yarahmadi S, Khodakarim S, Mahdavi Hezaveh A, et al. The patterns of Non-communicable disease Multimorbidity in Iran: A Multilevel Analysis. Sci Rep 2020 Feb 20;10(1):3034 [FREE Full text] [doi: 10.1038/s41598-020-59668-y] [Medline: 32080215]
13. Yao S, Cao G, Han L, Chen Z, Huang Z, Gong P, et al. Prevalence and Patterns of Multimorbidity in a Nationally Representative Sample of Older Chinese: Results From the China Health and Retirement Longitudinal Study. J Gerontol A Biol Sci Med Sci 2020 Sep 25;75(10):1974-1980. [doi: 10.1093/gerona/glz185] [Medline: 31406983]
14. Zemedikun DT, Gray LJ, Khunti K, Davies MJ, Dhalwani NN. Patterns of Multimorbidity in Middle-Aged and Older Adults: An Analysis of the UK Biobank Data. Mayo Clin Proc 2018 Jul;93(7):857-866. [doi: 10.1016/j.mayocp.2018.02.012] [Medline: 29801777]
15. Roso-Llorach A, Violán C, Foguet-Boreu Q, Rodriguez-Blanco T, Pons-Vigués M, Pujol-Ribera E, et al. Comparative analysis of methods for identifying multimorbidity patterns: a study of 'real-world' data. BMJ Open 2018 Mar 22;8(3):e018986 [FREE Full text] [doi: 10.1136/bmjopen-2017-018986] [Medline: 29572393]
16. Ioakeim-Skoufa I, Poblador-Plou B, Carmona-Pírez J, Díez-Manglano J, Navickas R, Gimeno-Feliu LA, et al. Multimorbidity Patterns in the General Population: Results from the EpiChron Cohort Study. Int J Environ Res Public Health 2020 Jun 14;17(12):4242 [FREE Full text] [doi: 10.3390/ijerph17124242] [Medline: 32545876]
17. Fabbri E, Zoli M, Gonzalez-Freire M, Salive ME, Studenski SA, Ferrucci L. Aging and Multimorbidity: New Tasks, Priorities, and Frontiers for Integrated Gerontological and Clinical Research. J Am Med Dir Assoc 2015 Aug 01;16(8):640-647 [FREE Full text] [doi: 10.1016/j.jamda.2015.03.013] [Medline: 25958334]
18. Brunson JC, Laubenbacher RC. Applications of network analysis to routinely collected health care data: a systematic review. J Am Med Inform Assoc 2018 Feb 01;25(2):210-221 [FREE Full text] [doi: 10.1093/jamia/ocx052] [Medline: 29025116]
19. Hidalgo CA, Blumm N, Barabási AL, Christakis NA. A dynamic network approach for the study of human phenotypes. PLoS Comput Biol 2009 Apr;5(4):e1000353 [FREE Full text] [doi: 10.1371/journal.pcbi.1000353] [Medline: 19360091]
20. Amell A, Roso-Llorach A, Palomero L, Cuadras D, Galván-Femenía I, Serra-Musach J, et al. Disease networks identify specific conditions and pleiotropy influencing multimorbidity in the general population. Sci Rep 2018 Oct 29;8(1):15970 [FREE Full text] [doi: 10.1038/s41598-018-34361-3] [Medline: 30374096]
21. Kalgotra P, Sharda R, Croff JM. Examining health disparities by gender: A multimorbidity network analysis of electronic medical record. Int J Med Inform 2017 Dec;108:22-28. [doi: 10.1016/j.ijmedinf.2017.09.014] [Medline: 29132627]
22. Kalgotra P, Sharda R, Croff JM. Examining multimorbidity differences across racial groups: a network analysis of electronic medical records. Sci Rep 2020 Aug 11;10(1):13538 [FREE Full text] [doi: 10.1038/s41598-020-70470-8] [Medline: 32782346]
23. Jiang Y, Ma S, Shia B, Lee T. An Epidemiological Human Disease Network Derived from Disease Co-occurrence in Taiwan. Sci Rep 2018 Mar 14;8(1):4557 [FREE Full text] [doi: 10.1038/s41598-018-21779-y] [Medline: 29540710]
24. Westergaard D, Moseley P, Sørup FKH, Baldi P, Brunak S. Population-wide analysis of differences in disease progression patterns in men and women. Nat Commun 2019 Feb 08;10(1):666 [FREE Full text] [doi: 10.1038/s41467-019-08475-9] [Medline: 30737381]
25. Multimorbidity: a priority for global health research. The Academy of Medical Sciences. 2018. URL: https://acmedsci.ac.uk/file-download/99630838 [accessed 2020-10-12]
26. Chronic Condition Indicators for ICD-10-CM (beta version). Healthcare Cost and Utilization Project (HCUP). URL: https://www.hcup-us.ahrq.gov/toolssoftware/chronic_icd10/chronic_icd10.jsp [accessed 2020-10-12]
27. Menche J, Sharma A, Kitsak M, Ghiassian SD, Vidal M, Loscalzo J, et al. Disease networks. Uncovering disease-disease relationships through the incomplete interactome. Science 2015 Feb 20;347(6224):1257601 [FREE Full text] [doi: 10.1126/science.1257601] [Medline: 25700523]
28. Ahlgren P, Jarneving B, Rousseau R. Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient. J. Am. Soc. Inf. Sci 2003 Apr;54(6):550-560. [doi: 10.1002/asi.10242]
29. Egghe L, Leydesdorff L. The relation between Pearson's correlation coefficient r and Salton's cosine measure. J. Am. Soc. Inf. Sci 2009 May;60(5):1027-1036. [doi: 10.1002/asi.21009]
30. Freeman LC. Centrality in social networks conceptual clarification. Social Networks 1978 Jan;1(3):215-239. [doi: 10.1016/0378-8733(78)90021-7]
31. Brin S, Page L. The anatomy of a large-scale hypertextual Web search engine. Computer Networks and ISDN Systems 1998 Apr;30(1-7):107-117. [doi: 10.1016/s0169-7552(98)00110-x]
32. Martin T, Zhang X, Newman MEJ. Localization and centrality in networks. Phys Rev E Stat Nonlin Soft Matter Phys 2014 Nov;90(5-1):052808. [doi: 10.1103/PhysRevE.90.052808] [Medline: 25493835]
33. Verma A, Bang L, Miller JE, Zhang Y, Lee MTM, Zhang Y, DiscovEHR Collaboration. Human-Disease Phenotype Map Derived from PheWAS across 38,682 Individuals. Am J Hum Genet 2019 Jan 03;104(1):55-64 [FREE Full text] [doi: 10.1016/j.ajhg.2018.11.006] [Medline: 30598166]

34. Fortunato S, Barthélemy M. Resolution limit in community detection. Proc Natl Acad Sci U S A 2007 Jan 02;104(1):36-41 [FREE Full text] [doi: 10.1073/pnas.0605965104] [Medline: 17190818]
35. Blondel VD, Guillaume J, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. J. Stat. Mech 2008 Oct 09;2008(10):P10008. [doi: 10.1088/1742-5468/2008/10/p10008]
36. Hongsong C. Networks, Crowds, and Markets: Reasoning about a Highly Connected World (Easley, D. and Kleinberg, J.; 2010) [Book Review]. IEEE Technol. Soc. Mag 2013;32(3):10-30. [doi: 10.1109/mts.2013.2276667]
37. Bähler C, Huber CA, Brüngger B, Reich O. Multimorbidity, health care utilization and costs in an elderly community-dwelling population: a claims data based observational study. BMC Health Serv Res 2015 Jan 22;15:23 [FREE Full text] [doi: 10.1186/s12913-015-0698-2] [Medline: 25609174]
38. Kostev K, Jacob L. Multimorbidity and polypharmacy among elderly people followed in general practices in Germany. Eur J Intern Med 2018 Sep;55:66-68. [doi: 10.1016/j.ejim.2018.07.014] [Medline: 30025890]
39. Wong A, Boshuizen HC, Schellevis FG, Kommer GJ, Polder JJ. Longitudinal administrative data can be used to examine multimorbidity, provided false discoveries are controlled for. J Clin Epidemiol 2011 Oct;64(10):1109-1117. [doi: 10.1016/j.jclinepi.2010.12.011] [Medline: 21454049]
40. Violán C, Foguet-Boreu Q, Hermosilla-Pérez E, Valderas JM, Bolíbar B, Fàbregas-Escurriola M, et al. Comparison of the information provided by electronic health records data and a population health survey to estimate prevalence of selected health conditions and multimorbidity. BMC Public Health 2013 Mar 21;13:251 [FREE Full text] [doi: 10.1186/1471-2458-13-251] [Medline: 23517342]
41. McLean G, Gunn J, Wyke S, Guthrie B, Watt GCM, Blane DN, et al. The influence of socioeconomic deprivation on multimorbidity at different ages: a cross-sectional study. Br J Gen Pract 2014 Jul;64(624):e440-e447 [FREE Full text] [doi: 10.3399/bjgp14X680545] [Medline: 24982497]
42. Low LL, Kwan YH, Ko MSM, Yeam CT, Lee VSY, Tan WB, et al. Epidemiologic Characteristics of Multimorbidity and Sociodemographic Factors Associated With Multimorbidity in a Rapidly Aging Asian Country. JAMA Netw Open 2019 Nov 01;2(11):e1915245 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.15245] [Medline: 31722030]
43. Lenzi J, Avaldi VM, Rucci P, Pieri G, Fantini MP. Burden of multimorbidity in relation to age, gender and immigrant status: a cross-sectional study based on administrative data. BMJ Open 2016 Dec 21;6(12):e012812 [FREE Full text] [doi: 10.1136/bmjopen-2016-012812] [Medline: 28003289]
44. Schiøtz ML, Stockmarr A, Høst D, Glümer C, Frølich A. Social disparities in the prevalence of multimorbidity - A register-based population study. BMC Public Health 2017 May 10;17(1):422 [FREE Full text] [doi: 10.1186/s12889-017-4314-8] [Medline: 28486983]
45. Chua YP, Xie Y, Lee PSS, Lee ES. Definitions and Prevalence of Multimorbidity in Large Database Studies: A Scoping Review. Int J Environ Res Public Health 2021 Feb 09;18(4):1673 [FREE Full text] [doi: 10.3390/ijerph18041673] [Medline: 33572441]
46. Nicholson K, Almirall J, Fortin M. The measurement of multimorbidity. Health Psychol 2019 Sep;38(9):783-790. [doi: 10.1037/hea0000739] [Medline: 31021126]
47. Wang X, Yao S, Wang M, Cao G, Chen Z, Huang Z, et al. Multimorbidity among Two Million Adults in China. Int J Environ Res Public Health 2020 May 13;17(10):3395 [FREE Full text] [doi: 10.3390/ijerph17103395] [Medline: 32414117]
48. Bhavnani SK, Dang B, Penton R, Visweswaran S, Bassler KE, Chen T, et al. How High-Risk Comorbidities Co-Occur in Readmitted Patients With Hip Fracture: Big Data Visual Analytical Approach. JMIR Med Inform 2020 Oct 26;8(10):e13567 [FREE Full text] [doi: 10.2196/13567] [Medline: 33103657]
49. George M, Smith A, Sabesan S, Ranmuthugala G. Physical Comorbidities and Their Relationship with Cancer Treatment and Its Outcomes in Older Adult Populations: Systematic Review. JMIR Cancer 2021 Oct 13;7(4):e26425 [FREE Full text] [doi: 10.2196/26425] [Medline: 34643534]
50. Valderas JM, Starfield B, Sibbald B, Salisbury C, Roland M. Defining comorbidity: implications for understanding health and health services. Ann Fam Med 2009;7(4):357-363 [FREE Full text] [doi: 10.1370/afm.983] [Medline: 19597174]
51. Han X, Hou C, Yang H, Chen W, Ying Z, Hu Y, et al. Disease trajectories and mortality among individuals diagnosed with depression: a community-based cohort study in UK Biobank. Mol Psychiatry 2021 Nov;26(11):6736-6746 [FREE Full text] [doi: 10.1038/s41380-021-01170-6] [Medline: 34035478]
52. Khan A, Uddin S, Srinivasan U. Comorbidity network for chronic disease: A novel approach to understand type 2 diabetes progression. Int J Med Inform 2018 Jul;115:1-9. [doi: 10.1016/j.ijmedinf.2018.04.001] [Medline: 29779710]
53. Steffen A, Nübel J, Jacobi F, Bätzing J, Holstiege J. Mental and somatic comorbidity of depression: a comprehensive cross-sectional analysis of 202 diagnosis groups using German nationwide ambulatory claims data. BMC Psychiatry 2020 Mar 30;20(1):142 [FREE Full text] [doi: 10.1186/s12888-020-02546-8] [Medline: 32228541]
54. Corrigan P. How stigma interferes with mental health care. Am Psychol 2004 Oct;59(7):614-625. [doi: 10.1037/0003-066X.59.7.614] [Medline: 15491256]
55. Ivbijaro GO, Enum Y, Khan AA, Lam SS, Gabzdyl A. Collaborative care: models for treatment of patients with complex medical-psychiatric conditions. Curr Psychiatry Rep 2014 Nov;16(11):506 [FREE Full text] [doi: 10.1007/s11920-014-0506-4] [Medline: 25218604]

56. Smith SM, Wallace E, O'Dowd T, Fortin M. Interventions for improving outcomes in patients with multimorbidity in primary care and community settings. Cochrane Database Syst Rev 2021 Jan 15;1:CD006560 [FREE Full text] [doi: 10.1002/14651858.CD006560.pub4] [Medline: 33448337]
57. Jensen AB, Moseley PL, Oprea TI, Ellesøe SG, Eriksson R, Schmock H, et al. Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients. Nat Commun 2014 Jun 24;5:4022 [FREE Full text] [doi: 10.1038/ncomms5022] [Medline: 24959948]
58. Haug N, Deischinger C, Gyimesi M, Kautzky-Willer A, Thurner S, Klimek P. High-risk multimorbidity patterns on the road to cardiovascular mortality. BMC Med 2020 Mar 10;18(1):44 [FREE Full text] [doi: 10.1186/s12916-020-1508-1] [Medline: 32151252]
59. Regitz-Zagrosek V, Kararigas G. Mechanistic Pathways of Sex Differences in Cardiovascular Disease. Physiol Rev 2017 Jan;97(1):1-37 [FREE Full text] [doi: 10.1152/physrev.00021.2015] [Medline: 27807199]
60. Gerdts E, Regitz-Zagrosek V. Sex differences in cardiometabolic disorders. Nat Med 2019 Nov;25(11):1657-1666. [doi: 10.1038/s41591-019-0643-8] [Medline: 31700185]
61. Chun H, Yue Y, Wang Y, Dawa Z, Zhen P, La Q, et al. High prevalence of congenital heart disease at high altitudes in Tibet. Eur J Prev Cardiol 2019 May;26(7):756-759. [doi: 10.1177/2047487318812502] [Medline: 30419180]
62. Hinton RB, Ware SM. Heart Failure in Pediatric Patients With Congenital Heart Disease. Circ Res 2017 Mar 17;120(6):978-994. [doi: 10.1161/circresaha.116.308996]
63. Garcia AM, Beatty J, Nakano SJ. Heart failure in single right ventricle congenital heart disease: physiological and molecular considerations. Am J Physiol Heart Circ Physiol 2020 Apr 01;318(4):H947-H965 [FREE Full text] [doi: 10.1152/ajpheart.00518.2019] [Medline: 32108525]
64. Taylor AW, Price K, Gill TK, Adams R, Pilkington R, Carrangis N, et al. Multimorbidity - not just an older person's issue. Results from an Australian biomedical study. BMC Public Health 2010 Nov 22;10:718 [FREE Full text] [doi: 10.1186/1471-2458-10-718] [Medline: 21092218]
65. De Las Rivas J, Fontanillo C. Protein-protein interaction networks: unraveling the wiring of molecular machines within the cell. Brief Funct Genomics 2012 Nov;11(6):489-496. [doi: 10.1093/bfgp/els036] [Medline: 22908212]
66. van Oostrom SH, Gijsen R, Stirbu I, Korevaar JC, Schellevis FG, Picavet HSJ, et al. Time Trends in Prevalence of Chronic Diseases and Multimorbidity Not Only due to Aging: Data from General Practices and Health Surveys. PLoS One 2016;11(8):e0160264 [FREE Full text] [doi: 10.1371/journal.pone.0160264] [Medline: 27482903]
67. Yu W, Singh SS, Calhoun S, Zhang H, Zhao X, Yang F. Generalized anxiety disorder in urban China: Prevalence, awareness, and disease burden. J Affect Disord 2018 Jul;234:89-96. [doi: 10.1016/j.jad.2018.02.012] [Medline: 29524751]
68. Zulman DM, Asch SM, Martins SB, Kerr EA, Hoffman BB, Goldstein MK. Quality of care for patients with multiple chronic conditions: the role of comorbidity interrelatedness. J Gen Intern Med 2014 Mar;29(3):529-537 [FREE Full text] [doi: 10.1007/s11606-013-2616-9] [Medline: 24081443]
69. Piette JD, Kerr EA. The impact of comorbid chronic conditions on diabetes care. Diabetes Care 2006 Mar;29(3):725-731. [doi: 10.2337/diacare.29.03.06.dc05-2078] [Medline: 16505540]
70. Mercer SW, Gunn J, Bower P, Wyke S, Guthrie B. Managing patients with mental and physical multimorbidity. BMJ 2012 Sep 03;345:e5559. [doi: 10.1136/bmj.e5559] [Medline: 22945951]
71. Soley-Bori M, Bisquera A, Ashworth M, Wang Y, Durbaba S, Dodhia H, et al. Identifying multimorbidity clusters with the highest primary care use: 15 years of evidence from a multi-ethnic metropolitan population. Br J Gen Pract 2021 Oct 20:BJGP.2021.0325 [FREE Full text] [doi: 10.3399/BJGP.2021.0325] [Medline: 34782317]

Abbreviations

HDR: hospital discharge report
ICD-10: International Classification of Diseases, 10th revision
RR: relative risk
SCI: Salton cosine index

Edited by T Leung; submitted 13.01.21; peer-reviewed by K Nicholson, Y Zhu; comments to author 06.04.21; revised version received 06.05.21; accepted 12.01.22; published 25.02.22.

Please cite as:
Wang L, Qiu H, Luo L, Zhou L
Age- and Sex-Specific Differences in Multimorbidity Patterns and Temporal Trends on Assessing Hospital Discharge Records in Southwest China: Network-Based Study
J Med Internet Res 2022;24(2):e27146
URL: https://www.jmir.org/2022/2/e27146
doi:10.2196/27146
PMID:35212632

Original Paper

Acceptability of and Willingness to Take Digital Pills by Patients, the Public, and Health Care Professionals: Qualitative Content Analysis of a Large Online Survey

Astrid Chevance[1,2], MD, PhD; Axel Fortel[3], BSc; Adeline Jouannin[4,5,6], MSc, MD; Faustine Denis[7], MSc, MD; Marie-France Mamzer[6,8], MD, PhD; Philippe Ravaud[1,2,9], MD, PhD; Stephanie Sidorkiewicz[1,10], MD, PhD

[1]Center for Research in Epidemiology and Statistics, Université de Paris-French National Institute for Health and Medical Research, Paris, France

[2]Service d'épidémiologie clinique, Hôpital Hôtel Dieu, Assistance Publique- Hôpitaux de Paris, Paris, France

[3]Faculté de médecine de Créteil, Université Paris-Est-Créteil, Créteil, France

[4]Département de Médecine Générale, Université de Rennes, Rennes, France

[5]Centre d'investigation clinique de Rennes, Centre Hospitalo-Universitaire de Rennes, Université de Rennes-French National Institute for Health and Medical Research, Rennes, France

[6]Department of Ethics, Research, Translations, Centre de Recherche des Cordeliers, Sorbonne Université, Université de Paris, French National Institute for Health and Medical Research, Paris, France

[7]Department of Psychiatry, Groupe Hospitalo-Universitaire Paris Psychiatrie Neurosciences, Paris, France

[8]Medical ethics Unit, Hôpital Necker Enfants-Malades, Assistance Publique-Hôpitaux de Paris, Paris, France

[9]Department of Epidemiology, Columbia University Mailman School of Public Health, New York, NY, United States

[10]Département de Médecine Générale, Université de Paris-French National Institute for Health and Medical Research, Paris, France

Corresponding Author:
Astrid Chevance, MD, PhD
Center for Research in Epidemiology and Statistics
Université de Paris-French National Institute for Health and Medical Research
1, place du pravis de Notre Dame
Paris, 75004
France
Phone: 33 142348987
Email: astrid.chevance@gmail.com

Abstract

Background: Digital pills are pills combined with a sensor, which sends a signal to a patch connected to a smartphone when the pills are ingested. Health care professionals can access patient data from digital pills online via their own interface, thus allowing them to check whether a patient took the drug. Digital pills were developed for the stated goal of improving treatment adherence. The US Food and Drug Administration approved the first digital pills in November 2017, but the manufacturer withdrew its application to the European Medicines Agency in July 2020 because of insufficient evaluation.

Objective: As recommended for the evaluation of health technologies, this study assesses the prospective acceptability of and willingness to take digital pills among patients, the public, and health care professionals.

Methods: Participants were patients who were receiving long-term treatment for a chronic condition, public participants (both groups recruited from a representative sample), and health care professionals. Participants answered 5 open-ended questions regarding the acceptability of digital pills and 1 close-ended question regarding the willingness to take digital pills, which were developed in a preliminary qualitative study. We explored the 5 theoretical dimensions of acceptability by performing an abductive qualitative content analysis of all free-text responses. We assessed data saturation with mathematical models. We fitted a multivariate logistic regression model to identify the sociodemographic and health characteristics associated with the willingness to take digital pills.

Results: Between January 29, 2020, and April 18, 2020, 767 patients, 1238 public participants, and 246 health care professionals provided 11,451 free-text responses. We identified 98 codes related to the acceptability of digital pills: 29 codes on perceived clinical effectiveness (eg, sensor safety cited by 66/2251 participants, 29.5%), 6 on perceived burden (eg, increased doctors' workload, 164/2251 participants, 7.3%), 25 on perceived ethicality (eg, policing, 345/2251 participants, 15.3%), 30 codes on perceived opportunity (eg, exclusively negative perception, 690/2251 participants, 30.7%), and 8 on affective attitude (eg, anger,

541/2251, 24%). Overall, 271/767 (35.3%) patients, 376/1238 (30.4%) public participants, and 39/246 (15.8%) health care professionals reported willingness to take digital pills. This willingness was associated with male sex (odds ratio 1.98, 95% CI 1.62-2.43) and current use of a connected device to record health settings (with a dose–response relationship).

Conclusions: The prospective acceptability of and willingness to take digital pills were limited by clinical and ethical concerns both at the individual and societal level. Our results suggest that digital pills should not be considered a mere change in the form of drug administration but a complex intervention requiring specific evaluation before extended use in clinical routine practice as well as an ethical and legal framework to ensure safe and ethical collection and use of health data through a patient-centered approach.

(J Med Internet Res 2022;24(2):e25597) doi:10.2196/25597

KEYWORDS

acceptability; health technology assessment; clinical effectiveness research; ethics; digital pill; digital health; digital therapeutics; ingestible sensor; adherence

Introduction

More than one-third of people worldwide live with at least 2 chronic conditions; hence, an increasing number of people are receiving long-term treatment [1]. Control of many of these conditions depends on adherence to treatment, which can be defined as the degree to which the use of medication by the patient corresponds with the prescribed regimen [2]. However, about half of patients with chronic conditions do not take their medications as prescribed. Poor treatment adherence leads to a range of negative consequences such as poor control of disease, life-threatening events, treatment resistance, and increasing costs [3-8]. In 2003, the World Health Organization suggested that improving treatment adherence would have a greater impact than any improvement in specific medical treatments, and this has led to a continuous search for effective interventions to monitor and improve medication adherence (eg, blood drug dosage, self-reported patient diaries, smart pill dispensers) [2,9,10].

The most recent monitoring tool, approved in 2017 by the US Food and Drug Administration, is called a “digital pill.” Digital pills are pills combined with a sensor, which sends a signal to a patch connected to the smartphone of the patient when the pill is ingested. Health care professionals and caregivers can access patient data (eg, adherence, activity, heart rate) from digital pills online via their own interface, thus allowing them to check whether a patient took the drug. The sensor is an ingestible event marker made from a copper–magnesium pair of electrodes within a silicon insulating skirt disk 5 mm in diameter and 0.3 mm thick. On contact with gastric fluid, the sensor sends a unique digital code to allow identification and timestamping of the medication and dose form [11]. The first digital pill was an antipsychotic (aripiprazole) used for schizophrenia or bipolar disorder treatment, which was rapidly followed by a range of other digital pills for multiple medical conditions (eg, antihypertension, diabetes, viral diseases) [12].

Digital pills have been developed for the stated goal of improving medication adherence and optimizing treatment management; these are used to identify nonadherence to first-line therapies before more expensive second-line therapies can be considered [10,13]. The European Medicines Agency declared that the evaluation for the digital pill submitted was inadequate and withdrew the application on July 26, 2020; the quality of the available studies (observational studies or small trials) and the lack of clinical relevance of the outcomes (short-term outcomes related to adherence to digital pills, safety outcomes restricted to tolerance of the patch) were points of criticism that led to this withdrawal [12,14-17].

Moreover, guidelines for the evaluation of health technologies recommend taking into account patient and public perspectives, which has not been done yet for digital pills [18]. Acceptability is a multifaceted construct that reflects the extent to which people delivering or receiving a health care intervention consider it appropriate based on anticipated, present, or retrospective responses to the intervention [19]. It includes the extent to which the intervention is perceived likely to achieve its purpose (perceived clinical effectiveness), the amount of effort required to participate in the intervention (perceived burden), the extent to which the intervention aligns with individual values (perceived ethicality), the extent to which benefits/values must be compromised to engage in the intervention (perceived opportunity), and how people feel about the intervention (affective attitude) [19].

As part of new health technology assessment, this study aims at evaluating the prospective acceptability of and willingness to take digital pills by 3 subpopulations (patients, public participants, and health care professionals).

Methods

We conducted an online survey by recruiting a representative sample of the French general population, dichotomized into patients and public participants, and an additional convenience sample of health care professionals. The survey included 5 open-ended questions related to the acceptability of digital pills and 1 close-ended question regarding the willingness to take digital pills (Multimedia Appendices 1 and 2).

Participants and Recruitment

Patients were adults (≥18 years old) with at least one self-reported chronic condition for >6 months and who were prescribed long-term treatment for >1 month. Public participants included adults (≥18 years old) who did not meet the eligibility criteria to be classified as patients. Health care professionals were professionals prescribing, delivering, or monitoring pharmaceutical treatments (physicians, nurses, pharmacists, and

midwives). All participants were living in France, spoke French, and had access to the internet.

We used the 2 distinct methods to recruit participants:

- The representative sample was recruited using a quota sampling method in the IPSOS Access Panel [20]. Quotas were based on the known profile of the French population in terms of sex, age, socioprofessional category, and population density of the residential area. This sample was secondarily classified into patients and public participants according to the eligibility criteria. We calculated weights with the rim weighting method—raking (Multimedia Appendices 3) [21].
- The additional convenience sample of health care professionals was recruited via an online campaign by sending emails to professional associations and posting advertisements on social networks (Multimedia Appendices 4).

All participants provided informed electronic consent before enrollment. The Institutional Review Board CERAPHP.5, Paris, France, gave ethical approval (IRB0001198).

To recruit patients with various combinations of chronic conditions and socioeconomic status, perform quantitative analysis using logistic regression, and ensure that data saturation was reached based on previous experience with an online survey and considering the fact that 34.5% of the French population lives with a chronic condition, we targeted a representative sample of 2000 participants [22,23]. Regarding the convenience sample of health care professionals, we decided to stop recruitment after reaching 95% data saturation [22].

Survey Development and Data Collection

We conducted a preliminary qualitative study to draft the questions (wording, formatting, and content) of the survey and provide the investigators with a contextual framework for analyzing and interpreting textual data. A researcher trained in qualitative methods (AC) conducted face-to-face collaborative interviews with 5 patients, 5 public participants, and 3 health care professionals (Multimedia Appendix 5). From the results of this preliminary study, we developed an illustrated explanation of digital pills, 5 open-ended questions on the acceptability (Textbox 1) and 1 close-ended question on the willingness to take digital pills ("Would you agree to use this device for yourself?") (Multimedia Appendices 1 and 2).

Textbox 1. Open-ended questions to all participants on the acceptability of digital pills.

- What do you think of this device (transmitter + Bluetooth patch + smartphone app + access to data)? We want to know your immediate reaction: try to write down all the ideas that came to you when you first saw this device.
- What do you envisage are the positive aspects of this device (transmitter + Bluetooth patch + smartphone app + access to data)? For example, in what situation(s)/for whom could it be useful, what kind of benefit could it bring?
- What do you envisage are the negative aspects of this device (transmitter + Bluetooth patch + smartphone app + access to data)? For example, what drawbacks or risks do you foresee?
- If your doctor suggested a treatment using this device (transmitter + Bluetooth patch + smartphone app + access to data), what would your reaction be? What would you think of his/her approach?
- The manufacturer says that this device will allow for greater consistency between the prescription and the actual taking of medications. In your opinion, why would a person take their treatment more consistently if they were equipped with this device?

In addition, we collected sociodemographic data (eg, sex, age, level of education) for patients and public participants, professional status data (profession and duration of professional experience) for health care professionals, and health characteristics for all participants (eg, chronic condition, current use of a connected device to collect health data) (Multimedia Appendices 1 and 2).

The survey for patients and public participants was available online in a dedicated IPSOS platform. The survey for health care professionals was available in a LimeSurvey form. Both surveys were tested online for usability, clarity, and wording in a convenience sample of 10 patients, 20 public participants, and 5 health care professionals, whose feedback helped improve the final version.

Analyses

Acceptability of Digital Pills

We conducted a multiple-round qualitative content analysis of each free-text response by following a 3-step approach: inductive open coding of all units of meaning related to the acceptability of digital pills, inductive condensation of the codes, and inductive development of the themes and deductive classification of the themes in the 5 dimensions of acceptability [19,24].

In the first step, the team of coders (AC, AF, AJ, FD) read and reread the responses at the individual level (all responses to all questions individual by individual) and at the question level (all responses for all individuals question by question). The process of coding identified in each response units of meaning and assigned a code. A unit of meaning was defined as any expression in the text related to the acceptability of digital pills, as explained by Sekhon et al [19]. The coding was limited to the manifest content of the responses, except for the first open-ended question on the emotional reaction to the digital pill, for which the latent emotional content was interpreted when possible [25]. During a face-to-face meeting, all researchers coded together all responses for 100/2005 (5%) individuals of the representative sample and 50/246 (20%) individuals in the sample of health care professionals. Then, 1 researcher coded all responses (AC) and 3 others double-checked the codes (AF, AJ, FD). In case of discrepancies, consensus was reached by discussion with another researcher (SS). To reduce interpretation

bias, coders had different backgrounds and training (psychiatry, family medicine, training in clinical research, social science, cognitive psychology, and medical ethics).

The second and third steps occurred in a meeting in which the 4 coders and SS inductively condensed all the codes using semantic, psychological, and anthropological considerations. Then, they categorized the codes inductively into themes and classified them deductively into the 5 theoretical dimensions of acceptability [19].

We then calculated the number of citations for each code identified with the qualitative content analysis in each group using the weighted data for patients and public participants.

Data saturation was assessed for patients, public participants, and health care professionals with a mathematical model [22].

Willingness to Take Digital Pills

We calculated the number and proportion of participants who agreed to take digital pills. We estimated odds ratios (ORs) and 95% CIs using the weighted data of the representative sample to calculate the association of the willingness to take digital pills with all socioeconomic and health characteristics. We used the same data for a multivariate logistic regression model, including sex, age, socioprofessional status, residential area, chronic condition, current use of a connected device to collect health data, and ease to talk about treatment with the doctor. For the regression, we considered the response "I do not wish to answer" to be a refusal to take digital pills. The significance threshold for *P* values was calculated with Bonferroni correction to correct for multiple testing.

All quantitative analyses involved using the free software R v3.3 (R Foundation for Statistical Computing), and the package "survey."

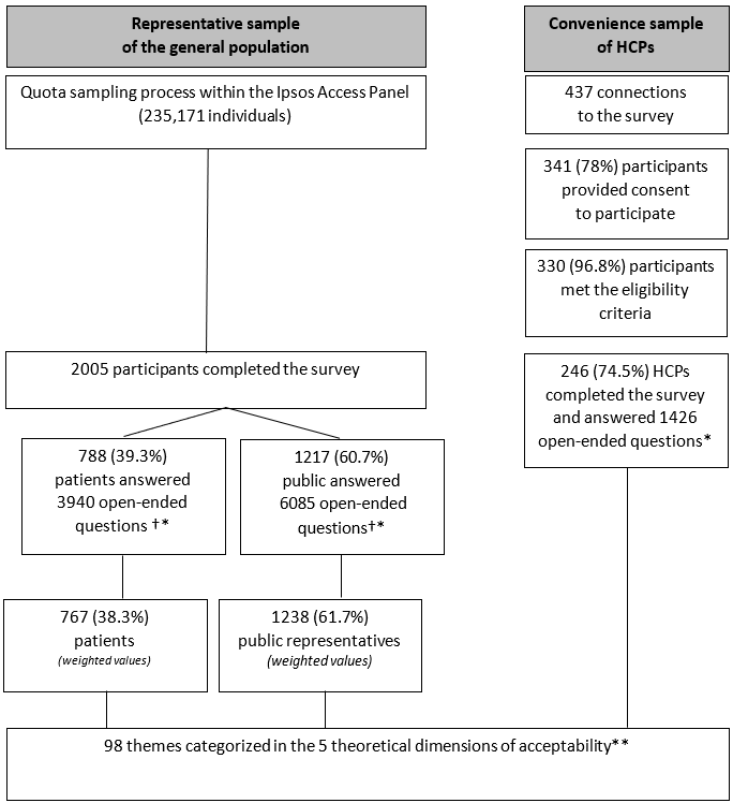
Data Sharing Statement

The deidentified quantitative data set will be shared on request (astrid.chevance@gmail.com).The textual data set will not be shared because of ethical restrictions.

Results

Between January 29, 2020, and February 5, 2020, IPSOS recruited a representative sample of 2005 participants (767 patients, 38.3%; 1238 public participants, 61.7%) (Figure 1). We recruited a convenience sample of 246 health care professionals between February 4, 2020, and April 18, 2020.

Figure 1. Flowchart. HCP: health care professional. †Unweighted numbers and proportions of patients and public representatives. *Participants answered 5 open-ended questions (panel 1). We added the number of open-ended questions answered for each type of participant. **Results presented in Figure 3 and Multimedia Appendix 10.



Participants

Among patients, 396/767 (51.6%) were women and the mean age was 54.6 (SD 15.6) years (Table 1). Hypertension was the most reported chronic condition (169/767, 22%), followed by diabetes (135/767, 17.6%) and chronic pain (81/767, 10.6%) (Multimedia Appendix 6).

Table 1. Sociodemographic and health characteristics of the representative sample and subgroups of patients and public participants.

Sociodemographic and health characteristics	Total (N=2005), n (%)	Patients[a] (n=767), n (%)	Public participants (n=1238), n (%)
Sex			
Female	1050 (52.4)	396 (51.6)	654 (52.8)
Male	955 (47.6)	371 (48.4)	584 (47.2)
Age (years)			
18-34	519 (25.9)	116 (15.1)	403 (32.6)
35-49	470 (23.4)	128 (16.7)	342 (27.6)
50-64	601 (30)	285 (37.2)	316 (25.5)
>65	415 (20.7)	238 (31)	177 (14.3)
Highest level of education			
Secondary school or under	75 (4.7)	43 (5.6)	52 (4.2)
Youth training	386 (19.3)	175(22.9)	210 (16.9)
High school graduate	471 (23.5)	172 (22.3)	299 (24.1)
Two-year university degree	458 (22.8)	174 (22.7)	284 (22.9)
Bachelor’s degree (BA, BS)	341 (17)	122 (15.9)	219 (17.7)
Master’s degree or beyond	254 (12.7)	80 (10.4)	174 (14.1)
Socioprofessional category			
Farmers	18 (0.9)	6 (0.8)	12 (1)
Self-employed professional workers	74 (3.7)	28 (3.6)	46 (3.7)
Senior managers	198 (9.9)	58 (7.6)	141 (11.4)
Technicians and associate professionals	305 (15.2)	68 (8.9)	236 (19.1)
Employees	349 (17.4)	89 (11.6)	260 (21)
Manual workers	265 (13.2)	76 (9.9)	188 (15.2)
Retired people	559 (27.9)	322 (42)	237 (19.1)
Unemployed	163 (8.1)	103 (13.4)	61 (4.9)
Student	74 (3.7)	17 (2.2)	57 (4.6)
Population density of the residential area (inhabitants)			
Rural city (<2000)	451 (22.5)	171 (22.3)	280 (22.6)
2000-19,999	360 (18)	138 (18)	211 (17)
20,000-99,999	276 (13.8)	109 (14.2)	161 (13)
≥100,000	605 (30.2)	229 (29.9)	377 (30.5)
Paris agglomeration	329 (16.4)	120 (15.6)	209 (16.9)
Number of visits to a doctor in the past year			
>10 times	208 (10.4)	140 (18.3)	68 (5.5)
5-10 times	590 (29.4)	336 (43.8)	254 (20.5)
<5 times	1046 (52.2)	286 (37.3)	759 (61.3)
Have not seen a doctor this year	152 (7.6)	5 (0.6)	148 (12)
Do not wish to answer	9 (0.4)	0 (0)	9 (0.7)
Ease to talk about the treatment with the doctor (eg, adherence, adverse events)			
Easy	1767 (88.1)	696 (90.8)	1071 (86.5)
Difficult	214 (10.7)	70 (9.1)	144 (11.6)
Do not wish to answer	24 (1.2)	1 (0.1)	23 (1.9)

Sociodemographic and health characteristics	Total (N=2005), n (%)	Patients[a] (n=767), n (%)	Public participants (n=1238), n (%)
Skipped the long-term treatment during the past month			
Never	N/A[b]	496 (64.7)	N/A
Once a week	N/A	198 (25.8)	N/A
Several times a week	N/A	54 (7)	N/A
Almost every day	N/A	15 (1.9)	N/A
Never started the prescribed treatment	N/A	3 (0.4)	N/A
Do not wish to answer	N/A	2 (0.2)	N/A
Checking health settings recorded by connected devices			
Daily	220 (11)	99 (12.9)	120 (9.7)
Weekly	213 (10.6)	85 (11.1)	128 (10.3)
Monthly	143 (7.2)	59 (7.7)	84 (6.8)
Rarely	257 (12.8)	99 (12.9)	159 (12.8)
Never	1168 (58.2)	425 (55.4)	743 (60)
Do not wish to answer	4 (0.2)	0 (0)	4 (0.3)
Agree to take a digital pill			
Yes	647 (32.3)	271 (35.3)	376 (30.4)
No	1261 (62.9)	461 (60.1)	800 (64.6)
Do not wish to answer	97 (4.8)	35 (4.6)	62 (5)

[a]Chronic condition was defined as at least 1 long-term treatment and 1 self-reported condition. The most common chronic conditions were hypertension, diabetes, chronic pain, thyroid disease, heart disease, asthma, and dyslipidemia. All chronic conditions are presented in Multimedia Appendix 6.

[b]N/A: not applicable.

A total of 99/767 (12.9%) patients and 120/1238 (9.7%) public participants reported checking health data recorded by a connected device daily, and 524/767 (68.3%) patients and 902/1238 (72.9%) public participants reported rarely checking or never checking data (Table 1).

Among health care professionals, there were 86/246 (35%) general practitioners, 50/246 (20.3%) psychiatrists, and 40/246 (16.3%) nurses (Table 2; Multimedia Appendix 7).

Table 2. Sociodemographic and health characteristics of the health care professionals (N=246).

Sociodemographic characteristics	Health care professionals, n (%)
Age (years), mean (SD)	35.5 (10.3)
Sex	
Female	155 (63)
Male	90 (36.6)
Other	1 (0.4)
Occupation	
Medical doctor	183 (74.4)
Nurse	40 (16.3)
Pharmacist	8 (3.2)
Midwife	9 (3.7)
Other	6 (2.4)
Medical specialists[a]	
General practitioner	86 (47)
Psychiatrist	50 (27.3)
Cardiologist	7 (3.8)
Other medical specialists[b]	40 (21.2)
Professional experience (years), mean (SD)	10.5 (10)
Has a chronic condition[c]	
Yes	28 (15.4)
No	218 (88.6)
Checking health settings recorded by the connected devices	
Every day	11 (4.5)
Several times per week	25 (10.2)
Several times per month	16 (6.5)
Rarely	60 (24.4)
Never	133 (54)
Do not wish to answer	1 (0.4)
Assessment of patient medication adherence	
Rarely	21 (8.5)
Depending on the patient and the medical situation	113 (46)
Systematically for each patient with long-term treatment and at each consultation	112 (45.5)
Agree to take a digital pill	
Yes	39 (15.8)
No	166 (67.5)
Do not wish to answer	11 (4.5)
Missing data	30 (12.2)

[a]n=183.

[b]For clarity, we report the results for only the 3 most prevalent medical specialties. The other medical specialists were anesthetists, intensive care unit physicians, emergency physicians, child psychiatrists, surgeons, obstetricians-gynecologists, hematologists, neurologist, gastroenterologists, dermatologists, endocrinologists, oncologists, pneumologists, infectious diseases specialists, rheumatologists, nephrologists, pediatric care specialists, palliative care specialists.

[c]Chronic condition was defined as at least 1 long-term treatment and 1 self-reported condition.

Acceptability of Digital Pills

The qualitative content analysis of the 11,451 open-text responses identified 98 codes related to acceptability. Data saturation was reached for patients (99.4%), public participants (100%), and health care professionals (97.7%) (Multimedia Appendix 8). We classified the 98 codes into the 5 theoretical dimensions of acceptability: perceived clinical effectiveness (29 codes, 29%), perceived burden (6 codes, 6%), perceived ethicality (25 codes, 25%), perceived opportunity (30 codes, 30%), and affective attitude (8 codes, 8%) [19]. We report the most-cited codes for each dimension with the proportion of citation and participants' quotes in Table 3. Figure 2 and Multimedia Appendix 9 report all 98 codes.

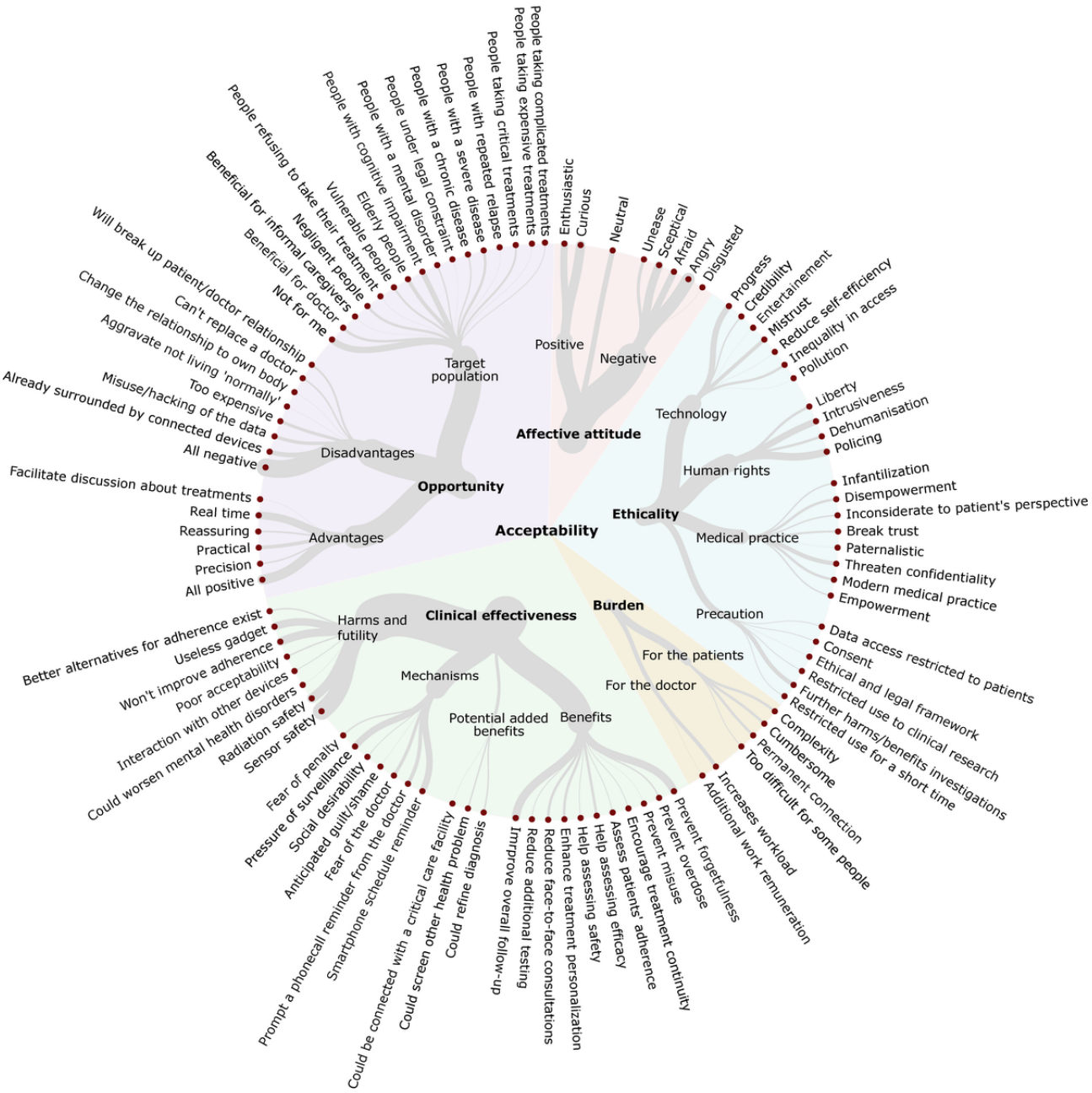
Table 3. Quotes of patients, public participants, and health care professionals that illustrate some of the codes regarding acceptability of digital pills.

Code	Patients (n=767)	Public participants (n=1238)	HCPs[a] (n=246)	Examples of quotes
Congruent with value of progress	59 (7.7)	131 (10.6)	29 (11.8)	• "I will try it because I have the culture of new technologies." (Public participant, man, 34 years old) • "I am for, because I believe in modernity and artificial intelligence." (Public participant, man, 43 years old) • "Only by accepting certain methods can one advance in medicine." (Nurse, woman, 46 years old)
Congruent with value of credibility	12 (1.6)	30 (2.4)	9 (3.7)	• "It's interesting, it's as close to reality as possible and probably the truth." (Public participant, man, 31 years old) • "There will no longer be any doubt about taking or forgetting to take medication, and the monitoring by the doctor that imposes honesty on the patient." (Public participant, man, 39 years old) • "This helps to avoid the medical errors that often occur when a patient lies about taking their medication." (Public participant, man, 29 years old)
Intrusiveness (conflict of value with privacy)	66 (8.6)	150 (12.1)	62 (25.2)	• "This is spying from inside my body. I would be worried about the side effects too." (Patient, woman, 64 years old) • "There is far too much indiscretion and invasion of privacy." (Public participant, woman, 55 years old) • "It's horrible! It feels as if it's really touching on intimacy. Ethically, it is frightening. That we use applications is one thing because we can limit its use in, in time, but a pill that we ingest is really to futuristic and too invasive." (Public participant, woman, 33 years old) • "It is a categorical refusal, I still prefer to take a drug that has less effect but I will never consciously swallow a capsule that allows other people to follow me or my lifestyle, even if already in society we are monitored in many areas without knowing it, my doctor has the right to suggest it after it is up to the person to accept it or not" (Patient, woman, 44 years old)
Dehumanization of patients	32 (4.2)	46 (3.7)	41 (16.7)	• "We're already guinea pigs and we're becoming robots... "(Patient, man, 64 years old) • "When you're sick, you're already dispossessed of your 'medical' life, especially in a hospital environment, so with that on top of it, no thanks." (Patient, woman, 53 years old) • "No thanks, I won't take any connected medication, I don't want to be kept under surveillance that much. It contributes to the dehumanisation of our society." (Neurologist, woman, 29 years old)
Inconsiderate to patient's perspective	25 (3.3)	21 (1.7)	30 (12.2)	• "Very bad idea, we're not in a dictatorship. People are free to take care of themselves or not. Even the sick are free." (Patient, man, 75 years old) • "Policing, (no) more freedom for the patient to stop a treatment that seems to be harming him without being observed." (Public participant, woman, 40 years old) • "A medication such as you're presenting it to me makes me think of forced treatment. If we refuse to take the medication or forget about it will our doctor contact us to explain ourselves?" (Patient, woman, 25 years old)

Code	Patients (n=767)	Public participants (n=1238)	HCPs[a] (n=246)	Examples of quotes
Improves overall follow-up	105 (13.7)	170 (13.7)	9 (3.7)	• "It allows us to monitor the effectiveness, which is already good! And also, that the doctor cares a little bit about his patient through closer monitoring." (Patient, man, 57 years old) • "The family doctor or specialist may monitor the illness on a day-to-day basis, whether good or bad, and may modify treatment or dosage if there are any problems, undesirable side effects or worsening of the patient's health. In addition, this device could also alert to new complications or illnesses that were not previously detected and that could be managed more quickly." (Public participant, man, 74 years old)
Poor acceptability	24 (3.1)	22 (1.8)	42 (17)	• "That's the problem: if a person doesn't follow a prescription, why would they want it to be known that they're doing whatever." (Patient, woman, 59 years old) • "In psychiatry, this process can be complicated for patients who are very often suspicious and persecuted, and all the more so if 'something enters their body' to keep them under surveillance... For the elderly, connected tools are not news except for the next generation." (Nurse, woman, 51 years old) • "I don't really see the point of such a complex system, when most of the information can be gathered through interrogation. The patients who are going to accept this device will probably be compliant patients and not the most problematic ones. Compliance is also a matter of education and not 'policing'. I don't see how being constantly kept under surveillance is going to get the patient to take their medication, other than by telling them they're going to get shouted at by their doctor, which is not our role." (Dermatologist, woman, 26 years old)
Useful for people with cognitive disorder	87 (11.3)	87 (7)	29 (11.8)	• "Unless I'm losing my mind, I wouldn't want to be kept under surveillance all the time." (Patient, woman, 77 years old) • "It might perhaps be useful for animals, but for humans, I don't see it. Unless, the person is not autonomous (e.g., senile dementia)." (Public participant, woman, 31 years old) • "For the elderly or people with amnesia: allows for better remote monitoring (and less cost for nurses to travel to the home)." (Patient, woman, 62 years old) • "I don't see the point in it for me, but for people who are out of their minds, why not? It's not very moral, but why not put it on for people without asking their opinion..." (Public participant, woman, 36 years old)

[a]HCPs: health care professionals.

Figure 2. Acceptability of connected drugs. Dots represent codes regarding acceptability of treatments identified by the qualitative content analysis of the responses for 767 patients, 1238 public participants, and 246 health care professionals. The width of the line is proportional to the number of spontaneous citations of the codes by the representative sample of patients and public participants (N=2005). The overarching categories in bold correspond to the 5 dimensions of the theoretical model of acceptability by Sekhon et al.



In Figure 2, which shows the acceptability of connected drugs, the dots represent codes regarding acceptability of treatments identified by the qualitative content analysis of the responses for 767 patients, 1238 public participants, and 246 health care professionals. The width of the line is proportional to the number of spontaneous citations of the codes by the representative sample of patients and public participants (N=2005). The overarching categories in bold correspond to the 5 dimensions of the theoretical model of acceptability by Sekhon et al [19]. *Perceived clinical effectiveness* is the extent to which the intervention is likely to achieve its purpose. *Perceived burden* is the amount of effort required to participate in the intervention. *Perceived ethicality* is the extent to which the intervention fits well with individual values. *Perceived opportunity* is the extent to which benefits/values must be given up to engage in the intervention. *Affective attitude* indicates how people feel about the intervention.

Perceived Clinical Effectiveness

Participants reported perceived benefits (11/98 codes, 11%) and harms (8/98 codes, 8%) and proposed further developments to the device to improve benefits (3/98 codes, 3%). Finally, they determined the underlying mechanisms of digital pills to improve medication adherence (7/98 codes, 7%).

Patients (105/767, 13.7%) and public participants (170/1238, 13.7%) believed that digital pills could improve follow-up. Potential harms cited were adverse events due to sensor ingestion (190/767 patients, 24.8%; 415/1238 public participants, 33.5%) or radiation from wireless technology

(102/767 patients, 13.3%; 212/1238 public participants, 17.1%). The underlying mechanisms cited to lead to a potential improvement in adherence included behavioral mechanisms such as a smartphone schedule reminder (96/767 patients, 12.5%; 122/1238 public participants, 9.9%) and the pressure of surveillance (77/767 patients, 10%; 141/1238 public participants, 11.4%).

Out of the 246 health care professionals, 64 (26%) considered digital pills to be useless gadgets, and 60 (24.4%) doubted the safety of the sensor. The mechanisms most cited to lead to improved adherence were the pressure of surveillance (73/246, 29.7%) and fear of arguing with the doctor (44/246, 17.9%).

Participants described the pressure of surveillance as a disciplinary mechanism based on social desirability, fear of being reprimanded, fear of social disqualification as a "bad patient," and fear of shame or guilt. Therefore, participants raised concerns about a policing patient–doctor relationship and the meaning of care in the future.

> *The fear of being kept under surveillance; it is being treated like a child, I find it, and even degrading.* [Patient, woman, 71 years old]

> *The person would take their treatment because they would know that the doctor has access to the actual taking. It is a social pressure: the doctor has particular expectations that the patient will want to respect.* [Patient, woman, 19 years old]

Perceived Burden

Participants distinguished doctors' burden (2/98 codes, 2%) from patients' burden (4/98 codes, 4%).

Patients and public participants mostly cited the complexity of the device (32/767, 4.2% and 46/ 1238, 3.7%, respectively) and the wearing of a cumbersome device (34/767, 4.4% and 41/1238, 3.3%, respectively) to be burdensome. Participants (62/767 patients, 8.1%; 65/1238 public participants, 5.3%) also reported the increased workload due to the processing of a large amount of data to be a problem.

Participants (12/767 patients, 1.6%; 14/1238 public participants, 1.1%; 15/246 health care professionals, 6.1%) identified contradictions in digital pills and found it unconvincing. Participants reported that taking medication seemed easier than applying a patch, activating the Bluetooth function, and ingesting the medication, especially for people who needed assistance with taking medication (eg, older people, individuals with cognitive impairment).

> *There will be better monitoring if the sick person is a mobile phone enthusiast: true for the young, less obvious for the elderly.* [Patient, man, 71 years old]

> *I find it hard to imagine someone who needs monitoring of their medication intake (dependent person, elderly, etc...) having such a patch device / smartphone.* [Patient, man, 36 years old]

Perceived Ethicality

Participants described whether taking a digital pill aligned with their values related to technology (7/98 codes, 7%), human rights (4/98 codes, 4%), and medical practice (8/98 codes). Participants also suggested precautions to limit the unethical use of digital pills (6/98 codes).

Regarding ethicality of the technology, patients and public participants mentioned that digital pills align with their values related to progress (59/767, 7.7% and 131/ 1238, 10.6%, respectively) and contradicted their values related to ecology (23/ 767, 3% and 44/1238, 3.5%, respectively). With respect to human rights, patients and public participants considered digital pills to be a policing tool (112/767, 14.6% and 143/1238, 11.6%, respectively) and worried about the associated intrusiveness (66/767, 8.6% and 150/1238, 12.1%, respectively). Regarding medical care, patients and public participants mostly reported concerns with confidentiality (44/767, 5.7% and 95/1238, 7.7%, respectively) and disempowerment of patients (44/767, 5.7% and 44/1238, 3.6%, respectively). To prevent the unethical use of digital pills, they requested further investigation of the associated harms and benefits (50/ 767 patients, 6.5%; 96/ 1238 public participants, 7.8%) and that the use of these pills be restricted to clinical research (9/767 patients, 1.2%; 11/1238 public participants, 0.9%).

In total, 29/246 (11.8%) health care professionals considered digital pills to be part of technological progress that aligned with their values; they mostly worried about policing (90/246, 36.6%) and intrusiveness (62/246, 25.2%).

Participants were afraid that digital pills might cause a breakdown in trust between patients and doctors (36/767 patients, 4.7%; 39/1238 public participants, 3.1%; 147/246 health care professionals, 59.8%), which is seen as the cornerstone of the patient–doctor relationship.

> *I don't like being kept under surveillance. There is a relationship of trust between my GP, my psychiatrist and me. This would undermine that trust because I would be policed.* [Patient, woman, 51 years old]

> *Ethically it's very questionable, it makes it look like we're 'spying on'/'policing' the patient. What about the patient's freedom? Could such a device really improve patient compliance? A good motivational interview seems more relevant. (It is better to act upstream = prevention and health promotion, therapeutic education, rather than downstream = bad compliance to be 'reprimanded') Not to mention that this would put the patient at odds vis-à-vis the doctor in case of poor compliance... And in this case, what about the doctor-patient trust relationship?* [General practitioner, woman, 29 years old]

Perceived Opportunity

Participants identified advantages (6/98 codes, 6%), disadvantages (8/98 codes, 8%), and specific target populations for whom they considered digital pills would be beneficial (16/98 codes, 16%).

A total of 135/767 (17.6%) patients and 200/1238 (16.2%) public participants had an overall positive perception of digital pills, and 213/767 (27.8%) patients and 408/1238 (33%) public participants had an overall negative perception of digital pills.

Patients and public participants mostly reported real-time data (70/767, 9.1% and 181/1238, 14.6%, respectively) as an advantage and the addition of another connected device in an already saturated environment (74/767, 9.6% and 86/1238, 6.9%, respectively) as a disadvantage. Regarding potential target populations, some highlighted that digital pills were not suitable for themselves (107/767 patients, 14%; 106/1238 public participants, 8.6%) and reported mostly people with cognitive impairment (87/767 patients, 11.3%; 87/1238 public participants, 7%) or older people (48/767 patients, 6.3%; 81/1238 public participants, 6.5%) to be suitable candidates.

Health care professionals pointed out the precision of digital pills (15/246, 6.1%) and the ease to discuss treatment (eg, tolerance, benefits, adherence) (11/246, 4.5%) as advantages and the cost (41/246, 16.7%) as a disadvantage. They identified people with cognitive impairment (29/246, 11.8%) and chronic conditions (19/246, 7.8%) as the main target population.

Participants identified people living under judicial constraint as a potential target population for digital pills (1/767 patients, 0.9%; 6/1238 public participants, 0.5%; 2/246 health care professionals, 4.6%); these people were depicted as dangerous, unwilling to undergo treatment, and already under surveillance and deprived of liberty, which—according to participants—would make the ethical regimen more flexible for them:

> *The person will take his treatment because he is compelled to do so by a court order. He is either a delinquent or a criminal. For someone who has not broken the law, there is no reason to proceed in this way.* [Public participant, man, 39 years old]

> *I think it can be a good idea for people who are sick (Alzheimer's for example) and forget to take their medication. Or for more complex cases such as people who are dangerous to society and the doctor needs to be sure that the person's treatment has been taken.* [Public participant, woman, 43 years old]

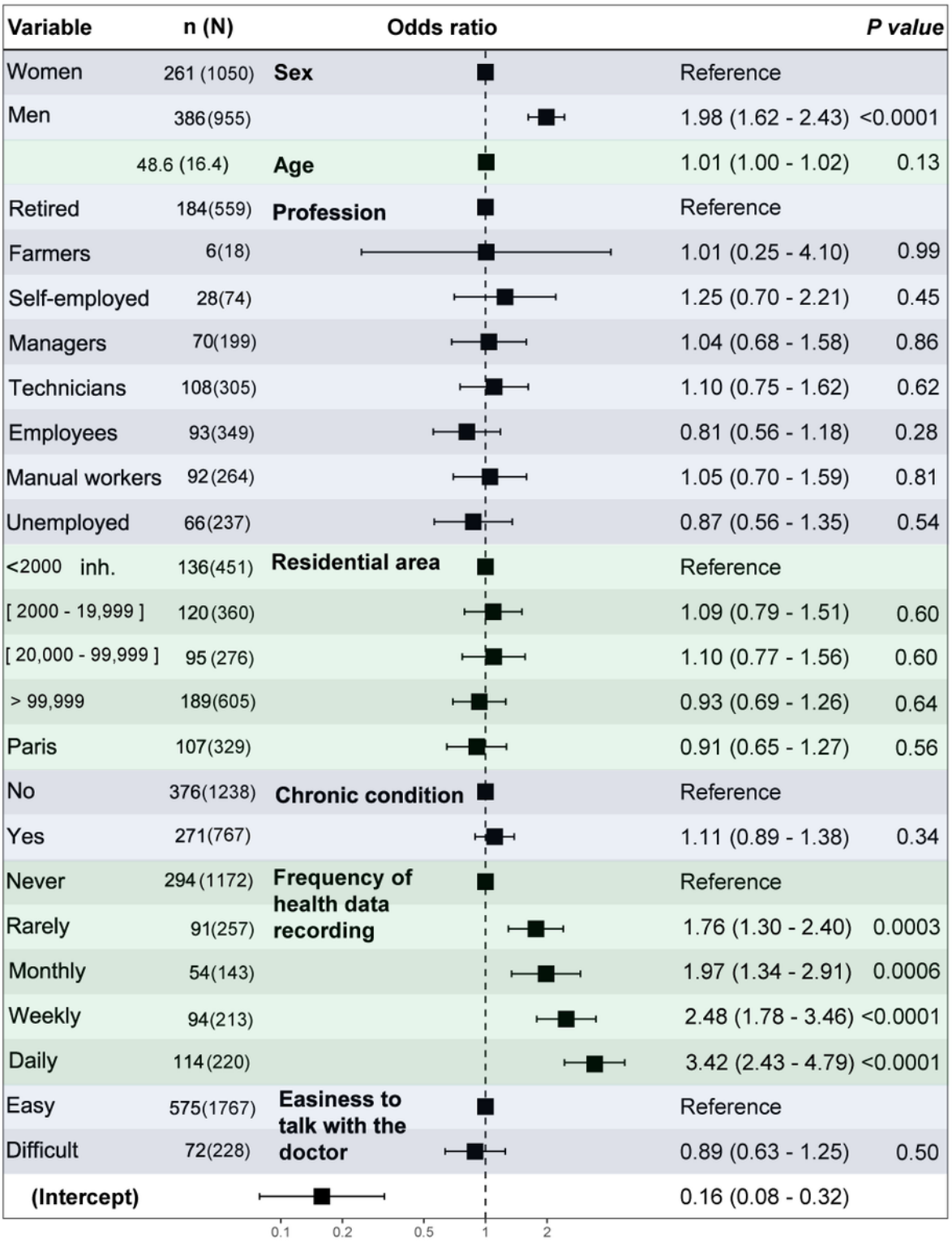
Affective Attitude

We identified 8 affective attitudes toward digital pills: enthusiasm (100/767 patients, 13%; 146/1238 public participants, 11.8%; 20/246 health care professionals, 8.1%), curiosity (141/767 patients, 18.4%; 170/1238 public participants, 13.7%; 31/246 health care professionals, 12.6%), balanced attitude (70/767 patients, 9.1%; 111/1238 public participants, 9%; 41/246 health care professionals, 16.7%), unease (32/767 patients, 4.2%; 70/1238 public participants, 5.7%; 7/246 health care professionals, 2.9%), skepticism (114/767 patients, 14.9%; 168/1238 public participants, 13.6%; 31/246 health care professionals, 5.3%); fear (56/767 patients, 7.3%; 175/1238 public participants, 14.1%; 13/246 health care professionals, 5.3%), anger (178/767 patients, 23.2%; 255/1238 public participants, 20.6%; 108/246 health care professionals, 43.9%), and disgust (7/767 patients, 0.9%; 4/1238 public participants, 0.3%; 1/246 health care professional, 0.4%). Because of lack of clarity and risk of overinterpretation, we could not code the affective attitude of 69/767 (9%) patients, 139/1238 (11.2%) public participants, and 12/246 health care professionals (4.9%). Compared to public participants, patients had more positive reactions (316/1238, 25.5% vs 241/767, 31.4%) and less negative reactions (672/1238, 54.3% vs 387/ 767, 50.5%). Compared to patients and public participants, health care professionals expressed more negative attitudes (142/246, 57.7%) and less positive attitudes (51/246, 20.7%).

Willingness to Take Digital Pills

In the representative sample, 647/2005 (32.3%) participants declared that they would use a digital pill: 271/767 (35.3%) patients, 376/1238 (30.4%) public participants, and 39/246 (15.8%) health care professionals. After adjustment, willingness to take digital pills by patients and public participants was significantly associated with male sex (OR 1.98, 95% CI 1.62-2.43) and integration of currently used connected devices to monitor health daily (OR 3.42, 95% CI 2.43-4.79) or rarely (OR 1.76, 95% CI 1.30-2.40), with a dose–response relationship (Figure 3). Complete analysis of other characteristics associated with the willingness to take a digital pill is presented with crude and adjusted ORs and 95% CIs in Multimedia Appendix 10. The forest plot (Figure 3) presents the ORs with their 95% CIs and the *P* value obtained with the logistic regression analysis of the association of the willingness to take connected drugs by sex, age, profession, population density of the residential area ("residential area"), chronic condition, frequency of health data recording with a connected device, and ease to talk about treatments with the doctor. The plot shows the number of participants per category (n) willing to take a connected drug among the total number of participants in the given category (N). The corrected threshold for statistical significance is *P*=.003. For the continuous variable "age," we present the mean age of the whole sample and the standard deviation.

Figure 3. Determinants of the willingness to take connected drugs among the representative sample of patients and public participants (N=2005). The second column present the number of people willing to take connected drugs (n) over the total number in the patient and public sample (N), except fot the variable "Age", presenting the mean and standard deviation (sd) of people willing to take connected drugs. inh.: inhabitants.



Discussion

Principal Findings

Overall, two-third of patients and public participants in this study refused to take digital pills. Their arguments were grounded in the fear of possible serious clinical and ethical harms. Health care professionals largely refused (177/246, 72%) to take digital pills citing disagreement with their professional deontology and ethics. Participants who accepted digital pills had positive a priori for technology considered as progress: people agreed to take it in the framework of research for a limited investigation time to help medical progress.

Participants who refused to take connected drugs, reported that these would be useful for other people, such as individuals with

cognitive impairment, individuals with chronic conditions or mental disorders, and those incarcerated, highlighting the view that potential clinical or ethical harms would be acceptable for these already vulnerable people. Given the condition of the designated populations, the intervention may be stigmatizing and may increase vulnerability.

To our knowledge, this study is the first to explore the acceptability of connected drugs in line with the recommendation that health technology assessment take into account patient and public perspectives [18]. We used a representative sample of the French population, thus allowing for generalizable results in the subsamples of patients and the public. In France, medication for chronic conditions is fully reimbursed, so when participants answered the questionnaire, they were not prematurely restricted by individual economic concerns but could detail their preferences. To limit the risk of capturing only a "hot reaction" and to mimic the process of a rational and balanced deliberation, we used 5 open-ended questions and asked the close-ended question about willingness to use connected drugs at the end of the survey. Only 16.7% of patients and public participants found no cons and 31% found no pros.

Limitations

The first limitation is the restriction to only 1 country as a consequence of the choice to recruit a representative sample. Because culture and health care system may affect the acceptability of digital pills, further studies are required in different settings.

A common limitation of using an online survey is the risk of selection bias because of the need for internet access. However, the target population for the digital pill must have access to smartphones and the internet. While the use of an online survey may have overestimated the acceptability of digital pills in the general population, it may not have done so in the target population. Moreover, to limit potential selection bias due to reading level, we created the questionnaire during a qualitative preliminary study and tested it for clarity.

We could have increased the reliability of the results by verifying the final codes and themes with the participants, but we minimized the risk of misinterpretation by limiting the condensation of codes to synonyms only, halting the inductive interpretation at a very basic level—which led to 98 codes in the final results—and using the participants' wording.

Another limitation is that the sample of health care professionals was not representative of the larger population of health care professionals.

Finally, we only investigated the prospective acceptability because digital pills were not available in France at the time of our study. Studies about concurrent acceptability (while taking connected drugs) and retrospective acceptability (after taking connected drugs) may be complementary and lead to different results.

Comparison With Prior Work

Our results are consistent with the concurrent and retrospective acceptability discussed in 18 observational studies and trials of digital pills, which involved small samples of 5 to 129 patients [26]. Moreover, some participants who refused digital pills for themselves found them useful for other people considered vulnerable (eg, people with cognitive or mental disorders or those incarcerated), which would lead to a risk of stigmatizing the use of digital pills.

This study added further ethical concerns to the well-known topics raised in the scientific literature (eg, consent, confidentiality, privacy) [12,26-30]. This shows the substantial importance of including patient and public perspectives in health technology assessment, as is already recommended but marginally done, with the risk of neglecting social, ethical, and political aspects of these technologies that are critical determinants of treatment effectiveness in real life settings [18]. A further strength is the use of a representative sample, which allows for generalizable results.

This survey contributes to advanced scientific and clinical reflections about medication adherence. As underlined by participants, digital pills seem to be a naïve solution considering the complexity of what determines medication adherence and nonadherence [12,31,32]. Digital pills appear to be burdensome and expensive reminders with an apparent contradiction—whether unintentional or intentional—between their complexity and the ability of those needing support for adherence. Particularly in the case of intentional nonadherence, digital pills would be either useless or efficient at the cost of intrusiveness, policing, and negative effects on the patient–doctor relationship.

Finally, digital pills could be useful to assess treatment efficacy in explanatory trials, but there are 2 prerequisites for extending their use in clinical practice. First, digital pills should be evaluated as a complex intervention according to specific standards and not as if they are a mere change in the form of drug administration [17,33]. Second, there is urgent need to develop an ethical and legal framework to ensure the safe and ethical collection and use of health data through a patient-centered approach [12,27,28].

Conclusion

Our results suggest that patients, the public, and health care professionals view connected drugs not as a promising new medical device to better monitor medication adherence but as a complex intervention with a possible impact on patient–doctor relationship and patient autonomy. The participants also raised additional concerns about burden of treatment, cost-effectiveness, and privacy, which need to be addressed in further investigations. Future studies should take into account the views of all stakeholders, ie, patients, potential prescribers, as well as health regulatory authorities and researchers, to ensure the safe and ethical collection and use of health data through a patient-centered approach.

Acknowledgments

We thank Florence Weber for the expert advice in social anthropology and Marion Plaze for expert advice in cognitive psychology. We thank Raphaël Porcher for statistical guidance, Elise Diard for expertise in the design of the figures, Nivantha Naidoo for help in interpreting the data, and Laura Smales for copyediting the manuscript. The study received a grant from the Fondation Université de Paris (19_DON22_001_U4). The funder had no role in the study design, data collection, data analysis, data interpretation, writing of the report, or decision to publish.

Authors' Contributions

AC is the guarantor of the study. AC, AJ, SS, PhR, and M-FM conceived and designed the analysis. AC, AJ, SS, and AF collected the data. AC, AJ, SS, AF, and FD performed the analysis. AC wrote the first draft. AC, SS, PhR, M-FM, FD, AF, and AJ edited the manuscript. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

The manuscript's guarantor (AC) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as originally planned (and, if relevant, registered) have been explained. All authors confirm that they had full access to all the data in the study and accept responsibility to submit for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Determinants of the willingness to take digital pills.
[DOCX File , 17 KB - jmir_v24i2e25597_app1.docx]

Multimedia Appendix 2
Questionnaire for healthcare professionals.
[DOCX File , 20 KB - jmir_v24i2e25597_app2.docx]

Multimedia Appendix 3
Calculation of the weight for the representative sample of the French general population (patients and public).
[DOCX File , 16 KB - jmir_v24i2e25597_app3.docx]

Multimedia Appendix 4
Methods for the online recruitment of healthcare professionals.
[DOCX File , 16 KB - jmir_v24i2e25597_app4.docx]

Multimedia Appendix 5
Preliminary qualitative study.
[DOCX File , 24 KB - jmir_v24i2e25597_app5.docx]

Multimedia Appendix 6
Chronic conditions of patients.
[DOCX File , 19 KB - jmir_v24i2e25597_app6.docx]

Multimedia Appendix 7
Description of healthcare professionals.
[DOCX File , 17 KB - jmir_v24i2e25597_app7.docx]

Multimedia Appendix 8
Data saturation.
[DOCX File , 18 KB - jmir_v24i2e25597_app8.docx]

Multimedia Appendix 9
Acceptability of digital pills.
[DOCX File , 71 KB - jmir_v24i2e25597_app9.docx]

Multimedia Appendix 10
Determinants of the willingness to take digital pills.
[DOCX File , 22 KB - jmir_v24i2e25597_app10.docx]

References

1. Chronic disease morbidity. OECDiLibrary. URL: https://www.oecd-ilibrary.org/sites/5101558b-en/index.html?itemId=/content/component/5101558b-en#figure-d1e16343 [accessed 2020-11-07]
2. Adherence to long-term therapies: Evidence for action. World Health Organization. 2003. URL: http://apps.who.int/iris/bitstream/handle/10665/42682/9241545992.pdf;jsessionid=4EDA2B0252381B3295C4756DFE01FB16?sequence=1 [accessed 2021-09-27]
3. Cutler RL, Fernandez-Llimos F, Frommer M, Benrimoj C, Garcia-Cardenas V. Economic impact of medication non-adherence by disease groups: a systematic review. BMJ Open 2018 Jan 21;8(1):e016982 [FREE Full text] [doi: 10.1136/bmjopen-2017-016982] [Medline: 29358417]
4. Bramley TJ, Gerbino PP, Nightengale BS, Frech-Tamas F. Relationship of blood pressure control to adherence with antihypertensive monotherapy in 13 managed care organizations. J Manag Care Pharm 2006 Apr;12(3):239-245. [doi: 10.18553/jmcp.2006.12.3.239] [Medline: 16623608]
5. Smith R. Adherence to antiretroviral HIV drugs: how many doses can you miss before resistance emerges? Proc Biol Sci 2006 Mar 07;273(1586):617-624 [FREE Full text] [doi: 10.1098/rspb.2005.3352] [Medline: 16537134]
6. Ho PM, Rumsfeld JS, Masoudi FA, McClure DL, Plomondon ME, Steiner JF, et al. Effect of medication nonadherence on hospitalization and mortality among patients with diabetes mellitus. Arch Intern Med 2006 Sep 25;166(17):1836-1841. [doi: 10.1001/archinte.166.17.1836] [Medline: 17000939]
7. Vanhove GF. Patient Compliance and Drug Failure in Protease Inhibitor Monotherapy. JAMA 1996 Dec 25;276(24):1955. [doi: 10.1001/jama.1996.03540240033024]
8. Yao X, Abraham NS, Alexander GC, Crown W, Montori VM, Sangaralingham LR, et al. Effect of Adherence to Oral Anticoagulants on Risk of Stroke and Major Bleeding Among Patients With Atrial Fibrillation. J Am Heart Assoc 2016 Feb 23;5(2):5:e003074 [FREE Full text] [doi: 10.1161/JAHA.115.003074] [Medline: 26908412]
9. Nieuwlaat R, Wilczynski N, Navarro T, Hobson N, Jeffery R, Keepanasseril A. Interventions for enhancing medication adherence. Cochrane Library. 2014 Nov 20. URL: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD000011.pub4/abstract?cookiesEnabled [accessed 2020-07-13]
10. Zijp TR, Mol PG, Touw DJ, van Boven JF. Smart Medication Adherence Monitoring in Clinical Drug Trials: A Prerequisite for Personalised Medicine? EClinicalMedicine 2019 Oct;15:3-4 [FREE Full text] [doi: 10.1016/j.eclinm.2019.08.013] [Medline: 31709407]
11. Plowman RS, Peters-Strickland T, Savage GM. Digital medicines: clinical review on the safety of tablets with sensors. Expert Opin Drug Saf 2018 Sep 13;17(9):849-852. [doi: 10.1080/14740338.2018.1508447] [Medline: 30073875]
12. Bittner B, Schmit Chiesi C, Kharawala S, Kaur G, Schmidt J. Connected drug delivery devices to complement drug treatments: potential to facilitate disease management in home setting. Med Devices (Auckl) 2019 Mar 12;12:101-127 [FREE Full text] [doi: 10.2147/MDER.S198943] [Medline: 30881151]
13. How does Proteus Discover Work. Proteus. URL: https://www.proteus.com/how-it-works/ [accessed 2020-05-28]
14. Cosgrove L, Cristea IA, Shaughnessy AF, Mintzes B, Naudet F. Digital aripiprazole or digital evergreening? A systematic review of the evidence and its dissemination in the scientific literature and in the media. BMJ Evid Based Med 2019 Dec 18;24(6):231-238. [doi: 10.1136/bmjebm-2019-111204] [Medline: 31320322]
15. Frias J, Virdi N, Raja P, Kim Y, Savage G, Osterberg L. Effectiveness of Digital Medicines to Improve Clinical Outcomes in Patients with Uncontrolled Hypertension and Type 2 Diabetes: Prospective, Open-Label, Cluster-Randomized Pilot Clinical Trial. J Med Internet Res 2017 Jul 11;19(7):e246 [FREE Full text] [doi: 10.2196/jmir.7833] [Medline: 28698169]
16. Osterberg L, Virdi N, Kim Y, Raja P, Cruz M, Savage G, et al. First Clinical Evaluation of a Digital Health Offering to Optimize Treatment In Patients With Uncontrolled Hypertension and Type 2 Diabetes. JACC 2016 Apr;67(13):2028. [doi: 10.1016/S0735-1097(16)32029-0]
17. Egilman AC, Ross JS. Digital medicine systems: an evergreening strategy or an advance in medication management? BMJ Evid Based Med 2019 Dec 21;24(6):203-204. [doi: 10.1136/bmjebm-2019-111265] [Medline: 31754072]
18. Gagnon M, Desmartis M, Lepage-Savary D, Gagnon J, St-Pierre M, Rhainds M, et al. Introducing patients' and the public's perspectives to health technology assessment: A systematic review of international experiences. Int J Technol Assess Health Care 2011 Jan 25;27(1):31-42. [doi: 10.1017/s0266462310001315]
19. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017 Jan 26;17(1):88 [FREE Full text] [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
20. Ipsos Online Access Panels. Ipsos. URL: https://www.ipsos.com/en-au/ipsos-online-access-panels [accessed 2021-09-23]
21. Kalton G, Flores-Cervantes I. Weighting methods. J Off Stat 2003;19(2):97 [FREE Full text]
22. Tran V, Porcher R, Tran V, Ravaud P. Predicting data saturation in qualitative surveys with mathematical models from ecological research. J Clin Epidemiol 2017 Feb;82:71-78.e2. [doi: 10.1016/j.jclinepi.2016.10.001] [Medline: 27789316]

23. Chevance A, Ravaud P, Tomlinson A, Le Berre C, Teufer B, Touboul S, et al. Identifying outcomes for depression that matter to patients, informal caregivers, and health-care professionals: qualitative content analysis of a large international online survey. Lancet Psychiatry 2020 Aug;7(8):692-702. [doi: 10.1016/S2215-0366(20)30191-7] [Medline: 32711710]
24. Graneheim UH, Lindgren B, Lundman B. Methodological challenges in qualitative content analysis: A discussion paper. Nurse Educ Today 2017 Sep;56:29-34. [doi: 10.1016/j.nedt.2017.06.002] [Medline: 28651100]
25. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov 01;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
26. Martani A, Geneviève LD, Poppe C, Casonato C, Wangmo T. Digital pills: a scoping review of the empirical literature and analysis of the ethical aspects. BMC Med Ethics 2020 Jan 08;21(1):3 [FREE Full text] [doi: 10.1186/s12910-019-0443-1] [Medline: 31914995]
27. Gerke S, Minssen T, Yu H, Cohen IG. Ethical and legal issues of ingestible electronic sensors. Nat Electron 2019 Aug 15;2(8):329-334. [doi: 10.1038/s41928-019-0290-6]
28. Véliz C. Three things digital ethics can learn from medical ethics. Nat Electron 2019 Aug 15;2(8):316-318. [doi: 10.1038/s41928-019-0294-2]
29. Holender A, Sutton S, De Simoni A. Opinions on the use of technology to improve tablet taking in >65-year-old patients on cardiovascular medications. J Int Med Res 2018 Jul 06;46(7):2754-2768 [FREE Full text] [doi: 10.1177/0300060518770578] [Medline: 29730949]
30. de Mendoza AH, Cabling ML, Dilawari A, Turner JW, Fernández N, Henderson A, et al. Providers' Perspectives on Adherence to Hormonal Therapy in Breast Cancer Survivors. Is there a Role for the Digital Health Feedback System? Health Technol (Berl) 2019 Mar 5;9(2):175-184 [FREE Full text] [doi: 10.1007/s12553-018-0267-x] [Medline: 31448187]
31. Sidorkiewicz S, Tran V, Cousyn C, Perrodeau E, Ravaud P. Development and validation of an instrument to assess treatment adherence for each individual drug taken by a patient. BMJ Open 2016 May 10;6(5):e010510 [FREE Full text] [doi: 10.1136/bmjopen-2015-010510] [Medline: 27165645]
32. Gadkari AS, McHorney CA. Unintentional non-adherence to chronic prescription medications: how unintentional is it really? BMC Health Serv Res 2012 Jun 14;12(1):98 [FREE Full text] [doi: 10.1186/1472-6963-12-98] [Medline: 22510235]
33. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008 Sep 29;337:a1655 [FREE Full text] [doi: 10.1136/bmj.a1655] [Medline: 18824488]

Abbreviations

OR: odds ratio

Edited by R Kukafka; submitted 08.11.20; peer-reviewed by K Bul, M Graber, S Sterckx; comments to author 31.01.21; revised version received 11.02.21; accepted 04.08.21; published 18.02.22.

Please cite as:
Chevance A, Fortel A, Jouannin A, Denis F, Mamzer MF, Ravaud P, Sidorkiewicz S
Acceptability of and Willingness to Take Digital Pills by Patients, the Public, and Health Care Professionals: Qualitative Content Analysis of a Large Online Survey
J Med Internet Res 2022;24(2):e25597
URL: https://www.jmir.org/2022/2/e25597
doi:10.2196/25597
PMID:35179509

Original Paper

The Efficacy of a Smartphone-Based App on Stress Reduction: Randomized Controlled Trial

Hyunchan Hwang[1], MSc, MD; Sun Mi Kim[1], MD, PhD; Bo Netterstrøm[2], MD, PhD, DrMedSci; Doug Hyun Han[1], MD, PhD

[1]Department of Psychiatry, College of Medicine, Chung-Ang University, Seoul, Republic of Korea

[2]Department of Occupational and Environmental Medicine, Bispebjerg University Hospital, Copenhagen, Denmark

Corresponding Author:
Doug Hyun Han, MD, PhD
Department of Psychiatry
College of Medicine
Chung-Ang University
102 Heukseok-ro, Dongjak-gu
Seoul, 156-756
Republic of Korea
Phone: 82 2 6299 1508
Email: hduk70@gmail.com

Abstract

Background: Stress management in the workplace is essential for a healthy mental and physical state. Due to technological advancements, individually tailored therapy and online cognitive behavioral therapy (CBT) are on the rise.

Objective: This study analyzed the efficacy of a smartphone app based on third-wave CBT tailored to an individual.

Methods: A randomized controlled trial was conducted with 126 participants who were divided into 2 groups. The intervention group used the smartphone app BetterLife for 10 weeks, while the control group was placed on a waiting list for the same duration. The Perceived Stress Scale–10 (PSS), Korean Utrecht Work Engagement Scale–9 (UWES), World Health Organization Quality of Life Assessment (WHOQOL), Beck Depression Inventory–II (BDI), and Beck Anxiety Inventory (BAI) were administered at baseline and after 10 weeks to both groups.

Results: Of the 126 participants, 11 dropped out during the trial. A 2-way repeated measure analysis of covariance was conducted, controlling for baseline BDI. There were greater improvements in PSS (F=24.33, P<.001, η^2=0.17) and UWESK scores (F=8.32, P=.0046, η^2=0.06) in the intervention group than in the control group. WHOQOL scores exhibited statistically significant improvement in the intervention group in the overall quality of life (F=8.19, P=.0049, η^2=0.06), physical health (F=8.87, P=.003, η^2=0.07), psychological health (F=13.32, P<.001, η^2=0.10), social relationships (F=19.43, P<.001, η^2=0.14), and environmental domains (F=10.14, P=.002, η^2=0.08) but not overall health (F=1.68, P=.20). BDI (F=7.17, P=.008, η^2=0.06) and BAI (F=6.00, P=.02, η^2=0.05) showed a statistically significant improvement in the intervention group, but this significance did not survive the Bonferroni correction (P<.005).

Conclusions: These results provide evidence that smartphone-based CBT is a viable option for reducing stress in the workplace.

Trial Registration: Clinical Research Information Service KCT0003231; https://cris.nih.go.kr/cris/search/detailSearch.do/15137

(J Med Internet Res 2022;24(2):e28703) doi:10.2196/28703

KEYWORDS

stress reduction; third-wave cognitive behavioral therapy; individual tailored treatment; randomized controlled trial; digital therapeutics

Introduction

Stress management is undoubtedly crucial to the mental health as well as physical health of individuals. Chronic and high stress have been linked to depression [1], anxiety [2], coronary heart disease [3], and increased mortality [4]. Stress can be acquired from all aspects of life. Nevertheless, work-related stress has been of interest to researchers, especially in South Korea, which

has one of the longest working hours within the Organization for Economic Cooperation and Development countries [5]. High stress within the workplace has also been associated with calculable economic burdens [6] and various long-term complications [7]. Therefore, stress management within the workplace is vital for both employees and employers.

Of the known methods for stress management, cognitive behavioral therapy (CBT) has been identified as effective [8,9]. Traditional CBT is performed by trained professionals such as psychiatrists or psychologists. It is usually conducted face to face or in a group setting. However, advancements in technology have allowed CBT to be computerized and administered via the internet, allowing it to be administered automatically with minimal guidance [8,10,11]. Such online CBTs have demonstrated efficacy in several areas, including the reduction of stress [12], depression, anxiety [13], and insomnia [14].

Although online CBT has advantages over traditional CBT in terms of anonymity and accessibility [15], in traditional CBT, the therapist can tailor the structured CBT to suit the client. This contrasts with most online CBTs, in which participants usually receive the same program that cannot be changed. Tailored online CBT has been implemented in treating depression and is effective, as shown in a meta-analysis [16]. However, Johansson et al [17] argued that tailored treatment is essential in addressing comorbidity, showing that tailored CBT may be more effective than ordinary internet CBT in treating depression. Stress is also associated with various disorders [1,2]. Thus, we developed an individualized CBT for stress management and performed a randomized controlled trial (RCT) to demonstrate its efficacy.

We hypothesized that using an online CBT-based app designed to help manage work stress for 10 weeks would result in a statistically significant improvement in stress-related scales compared to a waiting group.

Methods

Sample Size

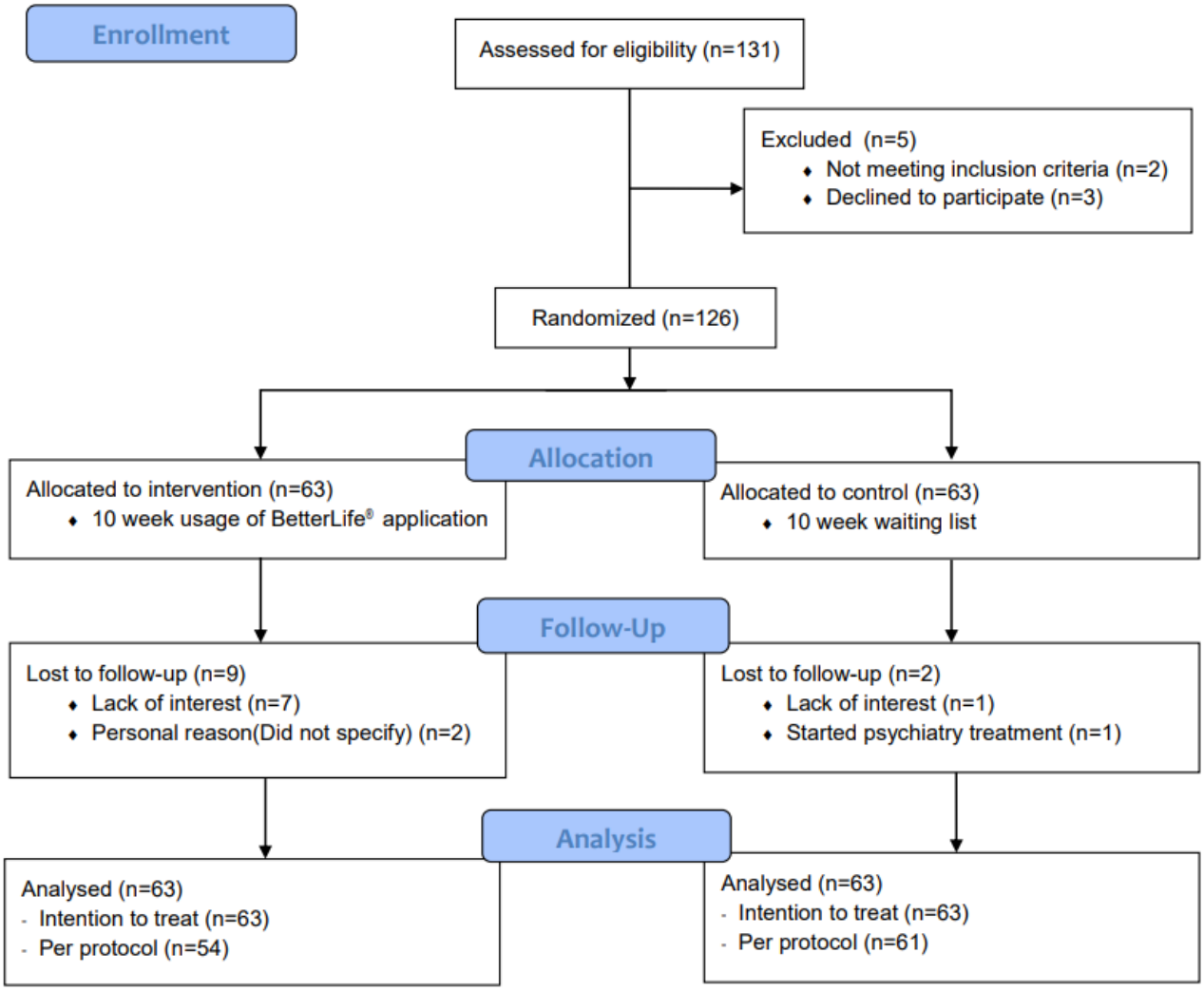
The sample size was calculated using G*Power (Heinrich-Heine-Universität Düsseldorf), based on a statistical test on repeated measure analysis of variance (ANOVA), since the covariance number was not determined. A type I error of .05 and statistical power of 0.8 were used. The correlation among repeated measures was set conservatively at 0, and based on previous similar studies [8], a mild to moderate effect size was estimated (effect size f=0.20). Based on these calculations, 51 participants were needed in each group, and after calculating a 20% dropout rate, the required total sample size was set at 128 participants.

Participants

Participants were recruited through billboard advertisements between November 2019 and January 2020. The inclusion criteria for the trial were as follows: (1) elevated perceived stress defined by a score of 14 or higher on the Perceived Stress Scale–10 (PSS); (2) indication that the reason for stress was mainly work-related (the reason for stress was considered to be work-related if the participant could identify one or more stress-related factors that they experienced in the work environment and stress in other areas such as family were minor in comparison to work-related stress); (3) employment of at least 20 hours per week and not self-employed; (4) aged 18 to 60 years; and (5) ability to provide informed consent to participate after being given information about the trial and other information that the participant must know to participate. The exclusion criteria were as follows: (1) inability to provide informed consent; (2) education level below the 9th grade; (3) history of congenital brain disorder, cerebral palsy, or other acquired brain injuries; (4) history of neurological disorders; (5) severe anxiety, depression, or psychotic disorder as assessed by the Korean Symptom Checklist–95; or (6) history of drug or alcohol abuse.

In total, 131 individuals who had work-related stress expressed interest in the trial. Of these, 3 withdrew participation after being informed of the detailed trial protocol; written consent was received from the remaining 128 participants. Two patients were excluded because of clinically relevant scores in the Korean Symptom Checklist–95. A total of 126 participants were included in the trial and randomized into 2 parallel groups with equal allocation ratios (Figure 1). Randomization was performed by an external researcher from a separate institute, and simple randomization via a random number generator using Excel (Microsoft Corp) was used. Each participant was assigned a random trial number at enrollment, and the external researcher was only provided with the individual trial number to ensure blinded randomization. The researcher performing data analysis (DHH) was also provided only with the individual trial number to blind the outcome assessment.

Figure 1. CONSORT 2010 flow diagram of enrollment.



The intervention group was given a random individual application ID to access the app, a link for downloading the app, and a manual of the program. Further, they were instructed to use the BetterLife program for 50 minutes per week for 10 weeks. They had the option to contact a designated person to ask for help in operating the app. Participants were sent a text message as a prompt to notify them of their app use every week during the trial. Members of the control group were placed on a waiting list for 10 weeks after which they were given access to the BetterLife program for the next 10 weeks. This was primarily done for ethical reasons, and data collected from the control group after using the BetterLife program were not used in this analysis. These data are being considered for use in secondary analyses, which can strengthen our research. This trial was approved by the institutional review board of Chung-Ang University Hospital (1712-008-304) and registered in the Clinical Research Information Service of Republic of Korea (KCT0003231), a member of the WHO International Clinical Trials Registry Platform. All participants enrolled in the trial provided written informed consent.

Of the 126 participants, 11 dropped out of the trial: 9 from the intervention group (14.3%) and 2 from the control group (3.2%). Of the 9 participants who were excluded from the intervention group, 2 left the trial due to personal reasons and 7 wished to leave due to lack of interest. Of the 2 participants who were excluded from the control group, 1 wished to leave the trial due to lack of interest and 1 was excluded because they were starting psychiatric treatment. Finally, 115 participants—54 in the intervention group and 61 in the control group—completed the 10-week trial and post-10-week data were collected. The 11 excluded participants' post-10-week data were produced by the baseline observation carried forward method, and data were analyzed using an intention-to-treat method. The average use time of the program was 56.40 (SD 8.42) minutes per week; this was calculated by using the program time of the 54 participants who finished the 10-week program only. No substantial harm or unintended effects were observed.

Data Collection

Participant demographic data were collected including age, sex, education level, marital status, alcohol consumption, and smoking habits. Information on employment was also collected, including company size, work field, type of employment, job grade, whether they handle direct customer complaints, income, workdays per week, work hours per day, work experience in the current work field, and job. Information regarding the number of late days, early leave days, and absent days of the

past month was also collected. The participants were asked to complete the PSS, Korean Utrecht Work Engagement Scale–9 (UWES), abbreviated version of the World Health Organization Quality of Life Assessment (WHOQOL), Beck Depression Inventory–II (BDI), and Beck Anxiety Inventory (BAI) to assess psychological and work stress-related status at baseline and after the 10-week intervention. PSS and the overall quality of life score of the WHOQOL at 10 weeks were set as the primary outcomes. Data were collected in a face-to-face setting by the researchers of Chung-Ang University Hospital at a separate location preferred by the participant. Participants received 100,000 Won (US $83) as compensation for their travel fees.

Perceived Stress Scale–10

The PSS was developed by Cohen et al [18] and validated in Korea by Park et al [19]. It is a widely used scale consisting of 10 items on a 5-point scale and is designed to measure the degree of stress experienced by individuals, with a high score indicating high perceived stress. Although the scale authors did not develop the scale with a cutoff value, the scores are often divided into 3 parts for clinical use: 0 to 13, low stress; 14 to 26, moderate stress; and >27, severe stress [20].

Korean Utrecht Work Engagement Scale–9

UWES is one of the most frequently used scales related to work engagement. It calculates work engagement by measuring vigor, dedication, and work absorption through a 7-point frequency rating scale ranging from 0 (never) to 6 (always). It was developed by Schaufeli et al [21] and has been validated in Korea by Kim et al [22].

World Health Organization Quality of Life Assessment, Abbreviated

WHOQOL was developed by the WHO as a measure to assess the quality of life (QoL) in individuals across different cultures [23]. It has been validated in Korea by Min et al [24]. The WHOQOL consists of 26 questions and uses a 5-point scale from 1 to 5, with a higher score indicating better QoL. One question pertains to overall perception of QoL (overall QoL); another pertains to overall perception of health (overall health). The remaining 24 questions were calculated to measure an individual's QoL in each of 4 domains: physical health, psychological health, social relationships, and environment. The raw scores were converted into transformed scores (0 to 100) using the developers' suggested method for easier comparison.

Beck Depression Inventory–II

The BDI was initially developed by Beck in 1961 and revised in 1979 to BDI-IA and finally to BDI-II in 1996 to accommodate the changes in the *Diagnostic and Statistical Manual of Mental Disorders, fourth edition* diagnosis for depressive disorders [25]. It is a self-report questionnaire with 21 items, and each item is rated on a 4-point scale from 0 to 3. It has been validated in many countries across the world, including Korea [26]. The BDI was designed to measure the depressive symptoms of an individual during the past 2 weeks.

Beck Anxiety Inventory

The BAI was developed by Beck in 1988 [27]. It is a self-report questionnaire with 21 items. Similar to the previous BDI, it rates each item from 0 to 3 on a 4-point scale. The BAI is designed to measure anxiety independent of depressive symptoms and was validated in Korea by Yook et al [28] in 1997.

Intervention

BetterLife is a smartphone-based program for the treatment of stress, depression, anxiety, and sleep disorders used for guided self-help therapy; it can be used for both prevention and treatment. It is designed for anonymous treatment, targeting individuals with mild or moderate symptoms. BetterLife uses recognized manuals for CBT and problem-solving therapy and consists of approximately 600 modules of 7 types: test modules, psychoeducational modules, cognitive exercises, practical exercises, diary modules, notification modules, and comment modules. The test modules included the World Health Organization Well-Being Index (WHO-5) [29], the Major Depression Inventory (MDI) [30,31], the General Anxiety Test–7 (GAD-7) [32], and a short version of the Copenhagen Psychosocial Questionnaire [33]; the latter is used for assessment of work-related stressors. Psychoeducational modules are e-learning modules that provide knowledge about illnesses, treatment methods, and background information. They have talking-head videos with an overlay of animated graphical visuals. Cognitive exercises are interactive exercises programmed in HTML5, providing opportunities for users to work with their thoughts and feelings. The interactivity makes it possible for the user to work with their data registered in a database on a server. Practical exercises include physical exercises, relaxation exercises such as meditation, and exercises away from home. Diary modules are used to map activities and the development of mental conditions over time. Data from the registrations in the diary modules are saved so the user can see them. Notification modules ensure the user's adherence to the treatment flows, and comment modules provide feedback to the user's treatment progression. Comment modules follow tests and exercises.

All treatment components in the program are based on evidence-based methods documented in scientific studies [11,34]. BetterLife is powered by an advanced treatment flow engine that enables individualized transdiagnostic treatment and continuous follow-up of treatment results carried out in a dialog form with the user. This dialog also includes an online dialog with an attached therapist and a chosen friend (helper) who can help the user understand the psychoeducational parts and conduct cognitive and practical exercises. The tests in the program are used both for reference to the individual treatment interventions in the program and for the continuous monitoring of treatment progression. If severe symptoms are detected, the user is referred for external psychological help in a clinic or hospital—in case of suicidal thoughts by call function directly to an acute clinic. The user sees graphical charts of the test result history to experience treatment progression. BetterLife has a toolbox with all tests and exercises in the program from where the user can choose to redo a test or exercise at any time. The

primary test and exercise results are always visible on the main page of the BetterLife program.

The chance to use the WHO-5 questionnaire was provided every week to tailor to each user. Thus, the user could follow the improvement over time, even in a graph. If the test value is low (meaning low well-being), the user is referred to the depression (MDI) and anxiety (GAD-7) modules. When acceptable values of these tests are obtained (showing a low degree of depression and anxiety symptoms), the user automatically returns to the other parts of the program. Another feature to personalize the program to the user is the user's own estimation of stressors in work and daily life. The user is guided to modules specifically concerning the problems pinpointed by the user. An overview of the program can be found in the supplements (Multimedia Appendices 1 and 2).

Statistical Analysis

All data were analyzed using SPSS (version 19.0, IBM Corp). The data presented here concern the intervention group compared to the waiting list group (controls) at baseline and follow-up. Data on the controls after using the BetterLife program and those on the program's long-term effects are not presented. Baseline information, including demographic and psychological scales, was compared using the chi-square test, Fisher exact test, and independent *t* test. Intervention outcomes were measured using work-related scales and psychological scores. Changes in late days, early leave days, and absent days in the past month were also collected to determine whether the app could have a tangible effect on work life.

For all participant data, univariate logistic regression analyses were conducted to determine whether variables of independent factors—such as age, gender ratio, education level, marital status, alcohol consumption, smoking, BDI scores, BAI scores, type of work field, workdays, work hours, and work experience—could explain a significant amount of variance in the dependent variable (high perceived stress) after considering all other variables. Statistically significant independent factors were evaluated using a multivariate logistic regression analysis using a stepwise forward conditional method. High perceived stress was defined as a PSS score of 27 or higher.

Using 2-way repeated-measures analysis of covariance (ANCOVA; split-plot ANCOVA) with time as a within-factor and group as a between-factor, controlling covariates determined from the logistic regression and the changes in PSS, UWES, WHOQOL, BDI, and BAI scores were calculated. Bonferroni correction was used to compensate for multiple comparisons, and α=.05/10=.005. P<.005 was considered statistically significant. The number of late days, early leave days, and absent days were analyzed separately by 2-way repeated measure ANOVA with time as a within-factor and group as a between-factor since they were not collected as part of the initial intervention outcome.

Results

Comparison of Demographic Characteristics and Psychological Scale Scores between the Intervention and Control Groups

There were no significant differences in age, gender ratio, education level, marital status, alcohol consumption, smoking, company size, work field, type of employment, job grade, direct customer complaint handling, workdays, work hours, work experience, and late days, early leave days, and absent days in the past month between the intervention and control groups (Table 1). At baseline, there were no significant differences in PSS, UWES, WHOQOL, BDI, and BAI scores (Table 2).

Table 1. Comparison of demographic characteristics of participants between intervention and control groups.

	Intervention (n=63)	Control (n=63)	Test statistic	*P* value
Demographic information				
Age (years), mean (SD)	38.6 (9.8)	37.3 (8.8)	0.75[a]	.46
Gender, n (%)	—[b]	—	2.45[c]	.12
Male	9 (14)	16 (25)	—	—
Female	54 (86)	47 (75)	—	—
Education, n (%)	—	—	0.89[c]	.64
High school	12 (19)	10 (16)	—	—
Undergraduate	46 (73)	45 (71)	—	—
Graduate	5 (8)	8 (13)	—	—
Marital state, n (%)	—	—	1.56[c]	.21
Single	26 (41)	33 (52)	—	—
Married	37 (59)	30 (48)	—	—
Alcohol, n (%)	—	—	0.05[c]	.82
Yes	51 (81)	52 (83)	—	—
No	12 (19)	11 (18)	—	—
Smoking, n (%)	—	—	2.49[c]	.12
Yes	3 (5)	8 (13)	—	—
No	60 (95)	55 (87)	—	—
Workplace information				
Company size (employee), n (%)	—	—	2.24[d]	.72
<10	5 (8)	5 (8)	—	—
10-29	2 (3)	6 (10)	—	—
30-99	11 (18)	11 (18)	—	—
100-299	5 (8)	4 (6)	—	—
≥300	40 (64)	37 (59)	—	—
Type of work field, n (%)	—	—	2.23[d]	.85
Sales and services	3 (5)	1 (2)	—	—
Technical	2 (3)	3 (5)	—	—
Office	16 (25)	21 (33)	—	—
Professional	23 (37)	22 (35)	—	—
Civil servant/teacher	5 (8)	4 (6)	—	—
Other[e]	14 (22)	12 (19)	—	—
Type of employment, n (%)	—	—	0.95[c]	.33
Regular position	51 (81)	55 (87)	—	—
Temporary position	12 (19)	8 (13)	—	—
Job grade, n (%)	—	—	3.27[c]	.66
Staff	2 (3)	1 (2)	—	—
Administrative manager	12 (19)	12 (19)	—	—
Assistant manager	10 (16)	14 (22)	—	—
General manager	12 (19)	17 (27)	—	—

	Intervention (n=63)	Control (n=63)	Test statistic	*P* value
Director and higher	18 (29)	13 (21)	—	—
No job grade	9 (14)	6 (10)	—	—
Customer complaints, n (%)	—	—	0.93[c]	.34
Yes	41 (65)	46 (73)	—	—
No	22 (35)	17 (27)	—	—
Income (KRW[f]), n (%)	—	—	3.23[d]	.53
<₩2 million	15 (24)	10 (16)	—	—
₩2-3 million	30 (48)	34 (54)	—	—
₩3-4 million	8 (13)	9 (14)	—	—
₩4-5 million	5 (8)	8 (13)	—	—
>₩5 million	5 (8)	2 (3)	—	—
Work days/week[g], mean (SD)	5.0 (0.3)	5.1 (0.2)	–0.72[a]	.47
Work hours/day[g], mean (SD)	8.2 (0.9)	8.3 (0.7)	–1.16[a]	.25
Work experience in current work field (months)[g], mean (SD)	133.2 (103.6)	128.0 (91.8)	0.30[a]	.77
Work experience in current job (months)[h], mean (SD)	107.3 (103.4)	91.2 (90.7)	0.92[a]	.36
Number of late days in past month, mean (SD)	0.7 (1.8)	0.8 (1.9)	–0.29[a]	.77
Number of early leave days in past month, mean (SD)	0.3 (0.6)	0.3 (0.9)	–0.34[a]	.73
Number of absent days in past month, mean (SD)	0.1 (0.2)	<0.1 (0.1)	1.01[a]	.31

[a]2-tailed *t* test.
[b]Not applicable.
[c]Chi-square test.
[d]Fisher exact test.
[e]Other includes miscellaneous positions in the company <9 employees, company management position >10 employees, etc.
[f]KRW: South Korean Won. 1198 Won=$1 USD.
[g]One value missing.
[h]Four values missing.

Table 2. Comparison of baseline work stress-related scale scores.

	Intervention (n=63)	Control (n=63)	Test statistic[a]	*P* value
PSS[b]	21.6 (5.9)	20.1 (3.8)	1.67	.10
UWES[c]	2.6 (0.8)	2.8 (0.8)	–1.41	.16
WHOQOL[d]				
Overall QoL[e]	3.0 (0.8)	3.1 (0.7)	–1.15	.25
Overall health	2.9 (0.9)	2.9 (0.8)	0.21	.84
Physical health	56.1 (13.8)	58.4 (12.3)	–0.99	.33
Psychological	52.5 (15.1)	57.3 (12.9)	–1.92	.06
Social relationship	55.1 (18.1)	61.0 (15.8)	–1.94	.06
Environmental	58.2 (14.6)	60.4 (12.6)	–0.88	.38
BDI[f]	17.7 (9.5)	15.3 (7.7)	1.53	.13
BAI[g]	13.8 (9.3)	11.0 (7.3)	1.89	.06

[a]2-tailed *t* test.

[b]PSS: Perceived Stress Scale.

[c]UWES: Utrecht Work Engagement Scale.

[d]WHOQOL: World Health Organization Quality of Life Scale.

[e]QoL: quality of life.

[f]BDI: Beck Depression Inventory.

[g]BAI: Beck Anxiety Inventory.

Comparison of Symptom Improvement Between the Intervention and Control Groups

In logistic regression analysis, only baseline BDI scores in all participants were positively correlated with high perceived stress (B=0.20, Exp(B)=1.22, *P*<.001; Table 3). A separate paired *t* test of the repeated measures of the control group showed no significant differences after the Bonferroni correction (Multimedia Appendix 3).

Table 3. Results of univariate and multivariate logistic regression analysis for high perceived stress.[a]

	B[b]	*P* value	Exp(B)[c]
Univariate			
Age	0.03	.91	1.00
Gender	–0.22	.72	0.80
Education	__[d]	.68	—
Marital state	–0.28	.59	0.76
Alcohol	0.34	.59	1.40
Smoking	0.46	.67	1.58
Company size	—	.89	—
Types of work field	—	.87	—
Type of employment	0.98	.10	2.67
Job grade	—	.86	—
Customer complaints	–0.06	.92	0.94
Income	—	.61	—
Work days/week	–21.03	>.99	0
Work hours/day	–0.40	.27	0.67
Work experience in current work field	<0.01	.53	1.00
Work experience in current job	<0.01	.61	1.00
BDI[e]	0.20	<.001	1.22
BAI[f]	0.14	<.001	1.15
Multivariate			
BDI	0.20	<.001	1.22

[a]Multivariate logistic regression analysis was conducted with BDI and BAI as independent variables.

[b]B: logistic regression coefficient.

[c]Exp(B): e to the power of B (odds ratio).

[d]Not applicable.

[e]BDI: Beck Depression Inventory–II.

[f]BAI: Beck Anxiety Inventory.

In the split-plot ANCOVA, there was homogeneity of variances (P>.05) and covariances (P>.001), as assessed by the Levene test of homogeneity of variances and Box M test, respectively, in all categories except for the number of late days, early leave days, and absent days in the past month. The Mauchly test of sphericity was ignored because there were only two groups for each factor. Controlling baseline BDI scores, the intervention group showed greater improvement in the changes in PSS (F=24.33, P<.001, η^2=0.17) and UWESK scores (F=8.32, P=.0046, η^2=0.06) compared to the control group (Figure 2). WHOQOL scores also demonstrated a statistically significant interaction between treatment groups and pretests/posttests in overall QoL (F=8.19, P=.0049, η^2=0.06) and physical health (F=8.87, P=.003, η^2=0.07), psychological (F=13.32, P<.001, η^2=0.10), social relationships (F=19.43, P<.001, η^2=0.14), and environmental domains (F=10.14, P=.002, η^2=0.08) but not overall health (F=1.68, P=.20; Table 4). Although BDI (F=7.17, P=.008) and BAI (F=6.00, P=.02) showed traditionally low P values, they did not survive the Bonferroni correction. Unadjusted ANOVA yielded similar results except for BDI (F=9.67, P=.002, η^2=0.07), which showed a significant interaction between treatment groups and pretests/posttests (Multimedia Appendix 4).

Figure 2. Repeated measure analysis of covariance controlling baseline Beck depressive inventory scores. Left: comparison of changes in Perceived Stress Scale (PSS) scores between intervention group and control group (F=24.33, P<.001). Right: comparison of changes in the Utrecht Work Engagement Scale–Korean Version (UWES) scores between intervention group and control group (F=8.32, P=.0046).

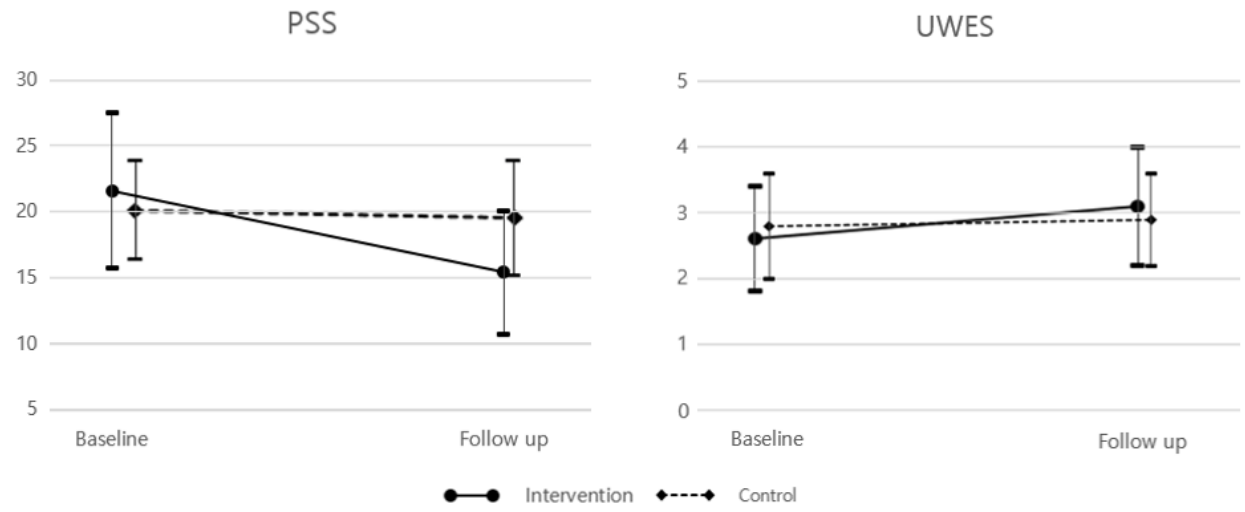


Table 4. Results of split-plot analysis of covariance for stress-related factors.

	Intervention (n=63)		Control (n=63)		Test statistics[a]		P value[b]
	Baseline	Follow-up	Baseline	Follow-up	F test[c]	η^2	
PSS[d]	21.6 (5.9)	15.4 (4.7)	20.1 (3.8)	19.6 (4.4)	24.33	0.17	<.001
UWESK[e] total	2.6 (0.8)	3.1 (0.9)	2.8 (0.8)	2.9 (0.7)	8.32	0.06	.0046
WHOQOL[f]							
Overall QoL[g]	3.0 (0.8)	3.6 (0.7)	3.1 (0.7)	3.3 (0.8)	8.19	0.06	.0049
Overall health	2.9 (0.9)	3.3 (0.8)	2.9 (0.8)	3.0 (0.9)	1.68	—[h]	.20
Physical health	56.1 (13.8)	65.8 (14.1)	58.4 (12.3)	60.7 (11.8)	8.87	0.07	.003
Psychological	52.5 (15.1)	63.6 (16.7)	57.3 (12.9)	59.5 (12.3)	13.32	0.10	<.001
Social relationship	55.1 (18.1)	66.6 (14.6)	61.0 (15.8)	59.1 (16.3)	19.43	0.14	<.001
Environmental	58.2 (14.6)	68.6 (12.5)	60.4 (12.6)	61.9 (10.9)	10.14	0.08	.002
BDI[i]	17.7 (9.5)	11.5 (9.2)	15.3 (7.7)	13.1 (7.6)	7.17	0.06	.008
BAI[j]	13.8 (9.3)	8.0 (8.7)	11.0 (7.3)	9.0 (7.0)	6.00	0.05	.02

[a]Statistics reported are for the interaction between intervention and time of each variable.

[b]P<.005 is considered significant.

[c]degree of freedom: 1123.

[d]PSS: Perceived Stress Scale.

[e]UWESK: Utrecht Work Engagement Scale.

[f]WHOQOL: World Health Organization Quality of Life Scale.

[g]QoL: quality of life.

[h]Not applicable.

[i]BDI: Beck Depression Inventory–II.

[j]BAI: Beck Anxiety Inventory.

Similarly, in the repeated measures ANOVA, homogeneity of variances (P>.05) and covariances (P>.001) were assessed by the Levene test of homogeneity of variances and Box M test, and the Mauchly test of sphericity was ignored. There were no significant interactions in the number of late days, early leave days, and absent days in the past month between the treatment groups and pretests/posttests (Table 5).

Table 5. Results of split-plot analysis of variance for the additional outcome.

	Intervention (n=63)		Control (n=63)		Test statistic[a,b]	*P* value
	Baseline	Follow-up	Baseline	Follow-up		
Number of late days in past month	0.7 (1.8)	0.4 (1.1)	0.8 (1.9)	0.7 (1.7)	0.48	.49
Number of early leave days in past month	0.3 (0.6)	0.1 (0.4)	0.3 (0.9)	0.3 (0.9)	0.74	.39
Number of absent days in past month	0.1 (0.2)	<0.1 (0.2)	<0.1 (0.1)	<0.1 (0.3)	0.49	.48

[a]Statistics (*F* test) reported are for the interaction between intervention and time of each variable.

[b]degree of freedom: 1124.

Sensitivity and Dropout Analyses

A sensitivity analysis was performed with the 54 intervention group and 61 control group participants who completed the trial. These data were evaluated separately using per-protocol analysis. Results showed a significant interaction between treatment groups and pretests/posttests in all scales except for the overall health of WHOQOL. Compared to the intention-to-treat analysis, the per-protocol analysis showed overall more significant (lower) *P* values and higher effect sizes (Multimedia Appendix 5). Little difference was found between the 2 analyses on the number of late days, early leave days, and absent days (Multimedia Appendix 6).

Dropout analysis was also performed, comparing the baseline values of the 11 participants who dropped out and the 115 participants who completed the trial. In the demographic data, work hours per day were significantly lower in the dropout group (t=3.63, P<.001; Multimedia Appendix 7). When comparing the baseline work-related stress and psychological scales, the dropout group showed significantly lower PSS scores (t=3.63, P<.001) and social relationship scores in the WHOQOL (t=–2.19, P=.03; Multimedia Appendix 8). The number of early leave days was also significantly lower in the dropout group (t=4.25, P<.001).

Discussion

Principal Findings

The BetterLife app, a program developed for the management of stress, effectively improved the degree of stress in people with work-related stress measured by PSS compared with control groups on a waiting list. It also improved work engagement as determined by the UWES. Furthermore, QoL improved in all domains except for overall health, according to the WHOQOL. However, the total number of late days, early leave days, and absent days showed no improvement in the 2 groups.

Intervention Effectiveness on Stress Reduction and Work Engagement

The effects of CBT on stress management in online settings have been well documented in previous studies [8,35]. A meta-analysis of web- and computer-based stress management interventions showed that these interventions effectively reduced stress and, on average, had a moderate effect size (Cohen d=0.43) on stress reduction [8]. However, recent studies have shown larger effect sizes. Asplund et al [36] identified a significant reduction in stress with a moderate to large effect size through an RCT with a guided internet-based stress management intervention. Ebert et al [10] also observed a large effect size in stress reduction using an internet- and mobile-based stress management program [10]. Our study showed that the intervention group had a significantly higher reduction in the PSS score from baseline to posttreatment when compared to the control group with a large effect size (F=24.83, P<.001, η^2=0.17). We estimate this large effect size to be because of meditation, emotional acceptance, and other third-wave CBT qualities incorporated in the program. Asplund et al [36] and Ebert et al [10] also used parts of third-wave CBT interventions, which could explain the large effect sizes of the results. Third-wave CBT helps participants to be more aware and accepting of their thoughts through ways such as meditation or behavior activation [37]. It has demonstrated strong efficacy in stress reduction [38].

The intervention group also showed improved work engagement, represented by improvements in the UWES. Work engagement, often considered the opposite of burnout, is a positive work-associated mindset with characteristics such as vigor, dedication, and absorption [39]. Extensive literature supports the positive benefits of work engagement. Halbesleben et al [40] demonstrated, through a study conducted on 587 employees in various occupations, that work engagement shared variance with job performance, implying that developing work engagement leads to positive outcomes in job performance and job retention. A meta-analysis of 7939 business units concluded that work engagement is related to meaningful business outcomes [41]. The authors of this study also implied that this improvement in employee work engagement and satisfaction could increase profits for businesses. Work engagement has also been known to affect absenteeism [42] and overall sickness absence [43]. However, our data did not show a significant difference in the number of sick days, early leave days, and absent days between baseline and postintervention. We estimate this to be because of the small number of sick, early leave, and absent days at baseline, subsequently requiring more statistical power than the trial's original design. However, several similar studies failed to demonstrate a significant improvement in sick leave days even with symptom improvement [44], which could imply that symptom improvement through CBT alone might not be enough for a tangible improvement in sick leave.

Intervention Effectiveness on Quality of Life

Improvement of QoL in all domains, including overall QoL, was seen in the intervention group. However, the perception of

overall health did not differ between the control and intervention groups. QoL improvement by CBT has been observed in literature: Hofmann et al [45] used a meta-analysis to observe the effect of CBT on QoL in anxiety disorders and found it to be effective. An RCT conducted using an internet-based self-guided stress management intervention for employees, similar to that in this study, also exhibited improvements in QoL [10], but only the mental health component of QoL improved (not the physical health component), which was similar to our findings. Although the physical domain of WHOQOL in our trial showed significant improvements, the effect size was the smallest among the 4 domains, and overall perception of health showed no significant improvement. This may be because improving physical health and related QoL is significantly more complex than improving mental health and related QoL, although possible. Therapies with a meditative component, such as yoga and mindfulness-based cognitive therapy, have been shown to improve health-related QoL [46]. CBT that improves sleep disturbance also induces an improvement in health-related QoL [47]. Furthermore, our program had separate modules dedicated to sleep improvement and meditation. A few participants gave positive feedback on the meditation module, which could have resulted in the mixed health-related QoL improvements. Future programs should emphasize this in CBT for further improvement in physical health.

The social relationship domain of QoL had the most significant effect size (η^2=0.14), in addition to our primary result, perceived stress. We attribute this improvement to the helper system that asked participants to find a helper, someone who can talk to and confide in about work. Participants also provided feedback that using this app with a helper allowed them to be more engaged in using the program and made the work environment more enjoyable. Improved social relationships within the workplace are estimated to increase employee well-being and company efficiency [48], constituting one reason for including CBT.

Changes in Depression and Anxiety Levels and Other Characteristics of the Program

Although our initial analysis showed that the BetterLife program significantly improved BDI and BAI levels, this significance did not survive the Bonferroni correction. We assume that this is because of the baseline observation carried forward method used to create the missing data of the participants who dropped out, which is one of the most conservative ways to estimate missing data. Notably, the per-protocol analysis showed that BDI and BAI scores were significantly improved using the BetterLife program with moderate effect size (F=10.15, P=.002, η^2=0.08 and F=8.21, P=.0049, η^2=0.07, respectively). The program incorporated modules that directly help deal with depression and anxiety, as controlling psychological symptoms is essential in reducing perceived stress. The efficacy of web-based CBTs has been proven in patients with major depressive disorder and anxiety disorder with mild and moderate symptoms [13]. Although the current intervention was aimed at employees with work stress, future trials could be performed on clinical groups to explore the program's effects.

The dropout rate of the intervention was 14.3%, slightly lower than the average dropout rate of internet-based CBTs according to a recent meta-analysis [49]. This, along with improved stress reduction, work engagement, and QoL, could result from the way the program was tailored for each participant, as described in the intervention section. The use of a helper combined with personalization as the user was guided through the program is a unique feature that makes BetterLife different from other internet-based programs dealing with stress. The feedback from the tests that participants took during the trial might also be a reason for the low degree of dropouts.

Although our study design did not include a direct comparison with face-to-face CBT, recent research comparing internet-based and face-to-face CBT showed similar results. Peter et al [14] showed similar efficacy in treating insomnia when comparing an online CBT with face-to-face CBT. A recent randomized noninferiority clinical trial showed online CBT to be noninferior to face-to-face CBT for health anxiety [50]. A systemic review and meta-analysis comparing online CBT and face-to-face CBT in psychiatric and somatic disorders reported that most disorders showed equivalent effects between the two groups [49]. These results emphasize the opinion that online CBT has become a viable and cost-effective option for face-to-face CBT, possibly because of advancements in technology and our familiarity with it. Unfortunately, we could not find any studies comparing online CBT with face-to-face CBT for stress reduction. Future studies are needed for stress management in this context.

Limitations

Our study has several limitations. First, only psychological scales were used in the trial. Although using objective measurements such as heart rate variance or blood cortisol levels would have complemented the results, we avoided this so that the participants would not have to visit a psychiatric clinic. Stigma regarding psychiatric interventions remains widespread in Korea, and having participants with work stress visit a psychiatric ward could negatively impact the trial [51]. Second, the average perceived stress was relatively high in both groups. Generalizing this result to a population with low perceived stress would be difficult with the current data. Third, most participants were female, and most had an education level of college diploma or higher. Although there are similar findings regarding demographics [12,46], this should be accounted for when applying the results to a broader population. Finally, only pre- and postintervention data were collected. Additional data from months after the intervention are needed to examine if the effects of the intervention persist over time.

Conclusions

BetterLife, a smartphone-based individually tailored CBT, effectively reduced stress and increased work engagement and QoL in people with work-related stress. This is a viable option for reducing workplace stress.

Acknowledgments

This study received funding from EUROSTARS-Project E! 11288 BETTERLIFE; the Innovation Fund Denmark under the Danish Ministry of Higher Education and Science; the South Korean Ministry of Trade, Industry, and Energy; and the Korea Institute for Advancement of Technology through the International Cooperative Research and Development Program.

Authors' Contributions

HH collected the data. HH and DHH analyzed the data. HH, SMK, BN, and DHH wrote the manuscript.

Conflicts of Interest

BN is a subcontractor for CONTEXT in development of the scientific part of BetterLife. CONTEXT will be the main contributor in sales and promotion of BetterLife. The other authors have no conflicts of interest to declare.

Multimedia Appendix 1
Screenshot of the BetterLife app.
[PNG File , 383 KB - jmir_v24i2e28703_app1.png]

Multimedia Appendix 2
Videorecording of a smartphone screen using the BetterLife app.
[MP4 File (MP4 Video), 1942 KB - jmir_v24i2e28703_app2.mp4]

Multimedia Appendix 3
Result of paired *t* test of baseline and follow-up of the control group.
[DOCX File , 26 KB - jmir_v24i2e28703_app3.docx]

Multimedia Appendix 4
Result of split-plot analysis of variance for stress-related factors (unadjusted).
[DOCX File , 23 KB - jmir_v24i2e28703_app4.docx]

Multimedia Appendix 5
Result of per-protocol analysis of split-plot analysis of covariance for stress-related factors.
[DOCX File , 23 KB - jmir_v24i2e28703_app5.docx]

Multimedia Appendix 6
Result of per-protocol analysis of split-plot analysis of variance for additional outcome.
[DOCX File , 24 KB - jmir_v24i2e28703_app6.docx]

Multimedia Appendix 7
Dropout analysis of demographic characteristics of participants.
[DOCX File , 31 KB - jmir_v24i2e28703_app7.docx]

Multimedia Appendix 8
Dropout analysis of baseline work stress-related information.
[DOCX File , 26 KB - jmir_v24i2e28703_app8.docx]

Multimedia Appendix 9
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1093 KB - jmir_v24i2e28703_app9.pdf]

References

1. Yang L, Zhao Y, Wang Y, Liu L, Zhang X, Li B, et al. The effects of psychological stress on depression. Curr Neuropharmacol 2015;13(4):494-504 [FREE Full text] [doi: 10.2174/1570159x1304150831150507] [Medline: 26412069]
2. Endler NS. Stress, anxiety and coping: the multidimensional interaction model. Canadian Psychology/Psychologie canadienne 1997;38(3):136-153. [doi: 10.1037/0708-5591.38.3.136]
3. Glozier N, Tofler GH, Colquhoun DM, Bunker SJ, Clarke DM, Hare DL, et al. Psychosocial risk factors for coronary heart disease. Med J Aust 2013 Aug 05;199(3):179-180. [doi: 10.5694/mja13.10440] [Medline: 23909539]

4. Steptoe A, Kivimäki M. Stress and cardiovascular disease. Nat Rev Cardiol 2012 Jun;9(6):360-370. [doi: 10.1038/nrcardio.2012.45] [Medline: 22473079]
5. OECD Factbook 2015-2016. URL: https://www.oecd-ilibrary.org/oecd-factbook-2015-2016_5jm5cxr8j2jc.pdf?itemId=%2Fcontent%2Fpublication%2Ffactbook-2015-en&mimeType=pdf [accessed 2022-01-31]
6. Hassard J, Teoh K, Cox T, Dewe P, Cosmar M, Gründler R, et al. Calculating the cost of work-related stress and psychosocial risks. 2014 Jun 27. URL: https://osha.europa.eu/sites/default/files/cost-of-work-related-stress.pdf [accessed 2022-01-31]
7. Nilsen C, Andel R, Fors S, Meinow B, Darin Mattsson A, Kåreholt I. Associations between work-related stress in late midlife, educational attainment, and serious health problems in old age: a longitudinal study with over 20 years of follow-up. BMC Public Health 2014 Aug 27;14:878 [FREE Full text] [doi: 10.1186/1471-2458-14-878] [Medline: 25159829]
8. Heber E, Ebert DD, Lehr D, Cuijpers P, Berking M, Nobis S, et al. The benefit of web- and computer-based interventions for stress: a systematic review and meta-analysis. J Med Internet Res 2017 Feb 17;19(2):e32 [FREE Full text] [doi: 10.2196/jmir.5774] [Medline: 28213341]
9. Hofmann SG, Asnaani A, Vonk IJJ, Sawyer AT, Fang A. The efficacy of cognitive behavioral therapy: a review of meta-analyses. Cognit Ther Res 2012 Oct 1;36(5):427-440 [FREE Full text] [doi: 10.1007/s10608-012-9476-1] [Medline: 23459093]
10. Ebert DD, Heber E, Berking M, Riper H, Cuijpers P, Funk B, et al. Self-guided internet-based and mobile-based stress management for employees: results of a randomised controlled trial. Occup Environ Med 2016 May;73(5):315-323. [doi: 10.1136/oemed-2015-103269] [Medline: 26884049]
11. Mathiasen K, Riper H, Ehlers LH, Valentin JB, Rosenberg NK. Internet-based CBT for social phobia and panic disorder in a specialised anxiety clinic in routine care: results of a pilot randomised controlled trial. Internet Interv 2016 May;4:92-98 [FREE Full text] [doi: 10.1016/j.invent.2016.03.001] [Medline: 30135794]
12. Heber E, Lehr D, Ebert DD, Berking M, Riper H. Web-based and mobile stress management intervention for employees: a randomized controlled trial. J Med Internet Res 2016;18(1):e21 [FREE Full text] [doi: 10.2196/jmir.5112] [Medline: 26818683]
13. Health Quality Ontario. Internet-delivered cognitive behavioural therapy for major depression and anxiety disorders: a health technology assessment. Ont Health Technol Assess Ser 2019;19(6):1-199 [FREE Full text] [Medline: 30873251]
14. Peter L, Reindl R, Zauter S, Hillemacher T, Richter K. Effectiveness of an online CBT-I intervention and a face-to-face treatment for shift work sleep disorder: a comparison of sleep diary data. Int J Environ Res Public Health 2019 Aug 24;16(17):1 [FREE Full text] [doi: 10.3390/ijerph16173081] [Medline: 31450619]
15. Andrews G, Basu A, Cuijpers P, Craske MG, McEvoy P, English CL, et al. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: an updated meta-analysis. J Anxiety Disord 2018 Apr;55:70-78 [FREE Full text] [doi: 10.1016/j.janxdis.2018.01.001] [Medline: 29422409]
16. Twomey C, O'Reilly G, Meyer B. Effectiveness of an individually-tailored computerised CBT programme (Deprexis) for depression: a meta-analysis. Psychiatry Res 2017 Dec;256:371-377. [doi: 10.1016/j.psychres.2017.06.081] [Medline: 28686935]
17. Johansson R, Sjöberg E, Sjögren M, Johnsson E, Carlbring P, Andersson T, et al. Tailored vs. standardized internet-based cognitive behavior therapy for depression and comorbid symptoms: a randomized controlled trial. PLoS One 2012;7(5):e36905 [FREE Full text] [doi: 10.1371/journal.pone.0036905] [Medline: 22615841]
18. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]
19. Park JO, Colvin KF. Psychometric properties of a Korean version of the Perceived Stress Scale (PSS) in a military sample. BMC Psychol 2019 Aug 30;7(1):611-629 [FREE Full text] [doi: 10.1186/s40359-019-0334-8] [Medline: 31470909]
20. Alharbi H, Alshehry A. Perceived stress and coping strategies among ICU nurses in government tertiary hospitals in Saudi Arabia: a cross-sectional study. Ann Saudi Med 2019;39(1):48-55 [FREE Full text] [doi: 10.5144/0256-4947.2019.48] [Medline: 30712051]
21. Schaufeli WB, Bakker A, Salanova M. The measurement of work engagement with a short questionnaire. Educ Psychol Meas 2016 Jul 02;66(4):701-716. [doi: 10.1177/00131644052824 71]
22. Ho Kim W, Park JG, Kwon B. Work engagement in South Korea. Psychol Rep 2017 Jul 14;120(3):561-578. [doi: 10.1177/0033294117697085] [Medline: 28558613]
23. WHOQOL Group. Development of the World Health Organization WHOQOL-BREF quality of life assessment. Psychol Med 1998 May;28(3):551-558. [doi: 10.1017/s0033291798006667] [Medline: 9626712]
24. Min SK, Kim KI, Lee CI, Jung YC, Suh SY, Kim DK. Development of the Korean versions of WHO Quality of Life scale and WHOQOL-BREF. Qual Life Res 2002 Oct;11(6):593-600. [doi: 10.1023/a:1016351406336] [Medline: 12206580]
25. Beck A, Steer R, Brown G. Manual for the Beck Depression Inventory-II. San Antonio: Psychological Corporation; 1996.
26. Sung H, Kim J, Park Y, Bai D, Lee S, Ahn H. A study on the reliability and the validity of the Korean version of the Beck Depression Inventory-II (BDI-II). J Korean Soc Biol Ther Psychiatry 2008;14(2):201-212.
27. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol 1988 Dec;56(6):893-897. [doi: 10.1037//0022-006x.56.6.893] [Medline: 3204199]

28. Yook S, Kim Z. A clinical study on the Korean version of Beck Anxiety Inventory: a comparative study of patient and non-patient. Korean J Clin Psychol 1997;16(1):185-197 [FREE Full text]
29. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 Well-Being Index: a systematic review of the literature. Psychother Psychosom 2015;84(3):167-176 [FREE Full text] [doi: 10.1159/000376585] [Medline: 25831962]
30. Bech P, Rasmussen NA, Olsen LR, Noerholm V, Abildgaard W. The sensitivity and specificity of the Major Depression Inventory, using the Present State Examination as the index of diagnostic validity. J Affect Disord 2001 Oct;66(2-3):159-164. [doi: 10.1016/s0165-0327(00)00309-8] [Medline: 11578668]
31. Olsen LR, Jensen DV, Noerholm V, Martiny K, Bech P. The internal and external validity of the Major Depression Inventory in measuring severity of depressive states. Psychol Med 2003 Mar;33(2):351-356. [doi: 10.1017/s0033291702006724] [Medline: 12622314]
32. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
33. Kristensen TS, Hannerz H, Høgh A, Borg V. The Copenhagen Psychosocial Questionnaire—a tool for the assessment and improvement of the psychosocial work environment. Scand J Work Environ Health 2005 Dec;31(6):438-449 [FREE Full text] [doi: 10.5271/sjweh.948] [Medline: 16425585]
34. Mathiasen K, Riper H, Andersen TE, Roessler KK. Guided internet-based cognitive behavioral therapy for adult depression and anxiety in routine secondary care: observational study. J Med Internet Res 2018 Nov 28;20(11):e10927 [FREE Full text] [doi: 10.2196/10927] [Medline: 30487118]
35. Netterstrøm B, Friebel L, Ladegaard Y. Effects of a multidisciplinary stress treatment programme on patient return to work rate and symptom reduction: results from a randomised, wait-list controlled trial. Psychother Psychosom 2013;82(3):177-186. [doi: 10.1159/000346369] [Medline: 23548852]
36. Persson AR, Dagöö J, Fjellström I, Niemi L, Hansson K, Zeraati F, et al. Internet-based stress management for distressed managers: results from a randomised controlled trial. Occup Environ Med 2017 Aug 30:105-113 [FREE Full text] [doi: 10.1136/oemed-2017-104458] [Medline: 28855344]
37. Hetrick SE, Cox GR, Witt KG, Bir JJ, Merry SN. Cognitive behavioural therapy (CBT), third-wave CBT and interpersonal therapy (IPT) based interventions for preventing depression in children and adolescents. Cochrane Database Syst Rev 2016 Aug 09:CD003380. [doi: 10.1002/14651858.CD003380.pub4] [Medline: 27501438]
38. Chiesa A, Serretti A. Mindfulness-based stress reduction for stress management in healthy people: a review and meta-analysis. J Altern Complement Med 2009 May;15(5):593-600. [doi: 10.1089/acm.2008.0495] [Medline: 19432513]
39. Schaufeli W, Salanova M, Gonzalez-Roma V, Bakker A. The measurement of engagement and burnout: a two sample confirmatory factor analytic approach. J Happiness Stud 2002;3:71-92. [doi: 10.1023/A:1015630930326]
40. Halbesleben JR, Wheeler AR. The relative roles of engagement and embeddedness in predicting job performance and intention to leave. Work Stress 2008 Jul;22(3):242-256. [doi: 10.1080/02678370802383962]
41. Harter JK, Schmidt FL, Hayes TL. Business-unit-level relationship between employee satisfaction, employee engagement, and business outcomes: a meta-analysis. J Appl Psychol 2002 Apr;87(2):268-279. [doi: 10.1037/0021-9010.87.2.268] [Medline: 12002955]
42. Magee C, Gordon R, Robinson L, Caputi P, Oades L. Workplace bullying and absenteeism: the mediating roles of poor health and work engagement. Hum Resource Manag J 2017 May 30;27(3):319-334. [doi: 10.1111/1748-8583.12156]
43. Rongen A, Robroek SJW, Schaufeli W, Burdorf A. The contribution of work engagement to self-perceived health, work ability, and sickness absence beyond health behaviors and work-related factors. J Occup Environ Med 2014 Aug;56(8):892-897. [doi: 10.1097/JOM.0000000000000196] [Medline: 25099418]
44. Salomonsson S, Santoft F, Lindsäter E, Ejeby K, Ljótsson B, Öst L, et al. Cognitive-behavioural therapy and return-to-work intervention for patients on sick leave due to common mental disorders: a randomised controlled trial. Occup Environ Med 2017 Dec;74(12):905-912. [doi: 10.1136/oemed-2017-104342] [Medline: 28756415]
45. Hofmann SG, Wu JQ, Boettcher H. Effect of cognitive-behavioral therapy for anxiety disorders on quality of life: a meta-analysis. J Consult Clin Psychol 2014 Jun;82(3):375-391 [FREE Full text] [doi: 10.1037/a0035491] [Medline: 24447006]
46. Grensman A, Acharya BD, Wändell P, Nilsson GH, Falkenberg T, Sundin, et al. Effect of traditional yoga, mindfulness-based cognitive therapy, and cognitive behavioral therapy, on health related quality of life: a randomized controlled trial on patients on sick leave because of burnout. BMC Complement Altern Med 2018 Mar 06;18(1):80 [FREE Full text] [doi: 10.1186/s12906-018-2141-9] [Medline: 29510704]
47. Brenes GA, Danhauer SC, Lyles MF, Anderson A, Miller ME. Effects of telephone-delivered cognitive-behavioral therapy and nondirective supportive therapy on sleep, health-related quality of life, and disability. Am J Geriatr Psychiatry 2016 Oct;24(10):846-854 [FREE Full text] [doi: 10.1016/j.jagp.2016.04.002] [Medline: 27421617]
48. Jutengren G, Jaldestad E, Dellve L, Eriksson A. The potential importance of social capital and job crafting for work engagement and job satisfaction among health-care employees. Int J Environ Res Public Health 2020 Jun 15;17(12):4272 [FREE Full text] [doi: 10.3390/ijerph17124272] [Medline: 32549346]

49. Carlbring P, Andersson G, Cuijpers P, Riper H, Hedman-Lagerlöf E. Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. Cogn Behav Ther 2018 Jan;47(1):1-18. [doi: 10.1080/16506073.2017.1401115] [Medline: 29215315]
50. Axelsson E, Andersson E, Ljótsson B, Björkander D, Hedman-Lagerlöf M, Hedman-Lagerlöf E. Effect of internet vs face-to-face cognitive behavior therapy for health anxiety: a randomized noninferiority clinical trial. JAMA Psychiatry 2020 Sep 01;77(9):915-924 [FREE Full text] [doi: 10.1001/jamapsychiatry.2020.0940] [Medline: 32401286]
51. Park J, Jeon M. The stigma of mental illness in Korea. J Korean Neuropsychiatr Assoc 2016 Nov 30;55(4):299-309. [doi: 10.4306/jknpa.2016.55.4.299]

Abbreviations

ANCOVA: analysis of covariance
ANOVA: analysis of variance
BAI: Beck Anxiety Inventory
BDI: Beck Depression Inventory–II
CBT: cognitive behavioral therapy
GAD-7: General Anxiety Test–7
MDI: Major Depression Inventory
PSS: Perceived Stress Scale–10
QoL: quality of life
RCT: randomized controlled trial
UWES: Korean Utrecht work engagement scale–9
WHO-5: World Health Organization Well-Being Index
WHOQOL: World Health Organization quality of life assessment, abbreviated

Edited by A Mavragani; submitted 11.03.21; peer-reviewed by G Won, A Baharav; comments to author 10.05.21; revised version received 12.06.21; accepted 29.12.21; published 15.02.22.

Please cite as:
Hwang H, Kim SM, Netterstrøm B, Han DH
The Efficacy of a Smartphone-Based App on Stress Reduction: Randomized Controlled Trial
J Med Internet Res 2022;24(2):e28703
URL: https://www.jmir.org/2022/2/e28703
doi:10.2196/28703
PMID:35166687

Letter to the Editor

Understanding the Impact of Social Media Information and Misinformation Producers on Health Information Seeking. Comment on "Health Information Seeking Behaviors on Social Media During the COVID-19 Pandemic Among American Social Networking Site Users: Survey Study"

Hunter Boudreau[1], BS; Nikhi Singh[1], BS; Carter J Boyd[2], MD, MBA

[1]School of Medicine, University of Alabama at Birmingham, Birmingham, AL, United States

[2]Hansjörg Wyss Department of Plastic Surgery, New York University Langone Health, New York, NY, United States

Corresponding Author:
Hunter Boudreau, BS
School of Medicine
University of Alabama at Birmingham
1670 University Blvd
Birmingham, AL, 35233
United States
Phone: 1 251 545 0115
Email: Hboud@uab.edu

Related Articles:

Comment on: https://www.jmir.org/2021/6/e29802

Comment in: https://www.jmir.org/2022/2/e31569/

(J Med Internet Res 2022;24(2):e31415) doi:10.2196/31415

KEYWORDS

social media; internet; communication; public health; COVID-19; usage; United States; information seeking; web-based health information; online health information; survey; mistrust; vaccination; misinformation

We congratulate Neely and colleagues on their recent work [1] describing the utilization of social media platforms as a source of information regarding the COVID-19 pandemic. The authors suggested that the majority of health information disseminated on social media was not fact-checked with a health care professional [1]. Furthermore, their results demonstrated that subjects following more credible scientific sources on social media were more likely to receive the COVID-19 vaccine [1]. These findings are corroborated by a recent study that concluded that there is a statistically significant relationship between disinformation regarding COVID-19 and lower vaccination rates [2]. However, both studies primarily focused on individual consumers of social media. While these studies are representative samples of the US population, they are unlikely to adequately describe the >200 million Twitter users and are likely subject to selection and recall bias from participants. While analysis of the consumer is revealing, understanding the publishers of information is of equal intrigue and utility.

An alternative methodology for addressing the investigative question proposed by Neely and associates would be to quantify content related to the COVID-19 vaccine and vaccination efforts and further classify this content as informed or misinformed. Data points could include the number of views and the frequency these posts receive subsequent dissemination. This approach would transition the focus from the consumers to the producers of this information. Studies of the aforementioned design come with their own set of limitations; however, we feel it is better suited to address the questions of the authors. Regardless of the study or methodology, social media platforms continue to grow, and health care professionals must recognize the potential effect they can have on social media.

Across social media platforms, it has been previously demonstrated that pro-vaccine individuals are more likely to reference credible sources than those from "antivaccine" groups [3]. Major social media platforms such as Facebook, Twitter, and Instagram have partnered with the World Health Organization in an attempt to target and flag misinformation

[3,4]. This served to counter misinformed COVID-19 and other health information on social media. Given social media's high availability and massive user base, there is a tremendous opportunity for physicians and health care organizations to interact with the American public through these virtual platforms. Establishing a stronger social media presence at both the systems (hospital, national governing medical body, academic center) and individual level is an underutilized opportunity for disseminating health information in an accurate manner. Most physicians (90%) have a presence on social media; however, it is unclear what advocacy impact these accounts have [5]. The introduction of a verification process for posts containing health information may have merit. Implementation of such a policy may increase consumer faith in factual health information, potentially enhancing public health advocacy in campaigns such as COVID-19 vaccinations. It appears that social media has a role to play in health care; an enhanced understanding of social media's scope of influence and increased physician representation may have a far-reaching impact.

Conflicts of Interest

None declared.

References

1. Neely S, Eldredge C, Sanders R. Health Information Seeking Behaviors on Social Media During the COVID-19 Pandemic Among American Social Networking Site Users: Survey Study. J Med Internet Res 2021 Jun 11;23(6):e29802 [FREE Full text] [doi: 10.2196/29802] [Medline: 34043526]
2. Wilson SL, Wiysonge C. Social media and vaccine hesitancy. BMJ Glob Health 2020 Oct;5(10) [FREE Full text] [doi: 10.1136/bmjgh-2020-004206] [Medline: 33097547]
3. Germani F, Biller-Andorno N. The anti-vaccination infodemic on social media: A behavioral analysis. PLoS One 2021;16(3):e0247642 [FREE Full text] [doi: 10.1371/journal.pone.0247642] [Medline: 33657152]
4. Gharib M. WHO Is Fighting False COVID Info On Social Media. How's That Going? NPR. 2021 Feb 09. URL: https://www.npr.org/sections/goatsandsoda/2021/02/09/963973675/who-is-fighting-false-covid-info-on-social-media-hows-that-going [accessed 2021-06-06]
5. Ventola CL. Social media and health care professionals: benefits, risks, and best practices. P T 2014 Jul;39(7):491-520 [FREE Full text] [Medline: 25083128]

Edited by T Leung; submitted 21.06.21; this is a non–peer-reviewed article;accepted 25.01.22; published 04.02.22.

Please cite as:
Boudreau H, Singh N, Boyd CJ
Understanding the Impact of Social Media Information and Misinformation Producers on Health Information Seeking. Comment on "Health Information Seeking Behaviors on Social Media During the COVID-19 Pandemic Among American Social Networking Site Users: Survey Study"
J Med Internet Res 2022;24(2):e31415
URL: https://www.jmir.org/2022/2/e31415
doi:10.2196/31415
PMID:35076408

Letter to the Editor

Digital Inclusion as a Foundation for Health Equity. Comment on "Expanding Video Consultation Services at Pace and Scale in Scotland During the COVID-19 Pandemic: National Mixed Methods Case Study"

Christopher Weatherburn[1,2], MBChB, MRCP, MRCGP, MSc

[1]National Health Service National Services Scotland, Edinburgh, United Kingdom

[2]National Health Service Tayside, Dundee, United Kingdom

Corresponding Author:
Christopher Weatherburn, MBChB, MRCP, MRCGP, MSc
National Health Service National Services Scotland
Gyle Square 1
South Gyle Crescent
Edinburgh, EH12 9EB
United Kingdom
Phone: 44 1382 436343 ext 6337
Email: christopher.weatherburn@nhs.scot

Related Article:

Comment on: https://www.jmir.org/2021/10/e31374

(J Med Internet Res 2022;24(2):e34702) doi:10.2196/34702

KEYWORDS

technology-enabled care; video consultations; quality improvement; COVID-19; PERCS framework; remote consultation; Scotland; general practice; digital inclusion; digital divide; digital health equity

This insightful paper by Wherton et al [1] suggests groundwork before the pandemic allowed health care services to rapidly extend the use of video consultations across Scotland. As a general practitioner and digital champion, working in a deprived area in a city in Scotland, I agree but strongly believe that front-line staff is fundamental to further increase use. The "asymmetric development, driven in some places by particular local enthusiasts" noted before the pandemic certainly continues, and additional local "champions" are key to further increased use.

The finding relating to the "focus on region-by-region" quality improvement approach noted can be explained by how the National Health Service is set up across Scotland, with 14 different health boards [2]. Although it is pertinent to point out the challenges of providing rural care, most citizens in Scotland live in nonrural areas [3].

My urban general practice embraced the opportunity to perform video consultations. We implemented these by deciding to "convert" telephone consultations to video consultations, when clinically appropriate. This occurred if the patient agreed and had the appropriate technology. We welcomed the improvement allowing the Near Me video service to send direct links to the virtual waiting room.

However, despite this improvement, when discussing video consultations with colleagues, we perceived that many patients then faced technical and accessibility issues. Patients frequently could not enter the virtual waiting room due to low levels of digital literacy.

A recent internal practice audit performed in August 2021 showed that only around half of the patients opted for a video consultation (n=67, 46%) if offered instead of a telephone consultation. Furthermore, of those who requested a video consultation, a large proportion was unsuccessful (n=31, 45%).

If a patient was unable to connect to the video consultation, they would not have the opportunity to complete the patient satisfaction survey, which is given out after the video consultation. Therefore, the patient survey, which reported that the majority of video consultations had no technical problems (n=18,817, 78%), is potentially an underestimate as it does not include some consultations in which patients were unable to access the service at all.

I am pleased to report that the Near Me video service has just been upgraded to include a Consult Now feature. This enables a one-time-only link to be sent, bringing the patient directly to the video call, without having to enter their details and access the virtual waiting room.

In an attempt to offer patients choice, I believe patients should be offered a video consultation instead of a telephone consultation, if they prefer. A video consultation may not objectively aid the consultation; however, it can assist with the clinical relationship.

To do this while addressing digital inclusion, including digital access support is fundamental. We must be mindful to continue to strive to reduce health inequalities and address the "digital inverse care law." I am personally very supportive of Scotland's digital health and care strategy, which aims to achieve world-leading levels of digital inclusion [4].

Conflicts of Interest

None declared.

Editorial Notice

The corresponding author of *"Expanding Video Consultation Services at Pace and Scale in Scotland During the COVID-19 Pandemic: National Mixed Methods Case Study"* declined to respond to this letter.

References

1. Wherton J, Greenhalgh T, Shaw S. Expanding Video Consultation Services at Pace and Scale in Scotland During the COVID-19 Pandemic: National Mixed Methods Case Study. J Med Internet Res 2021 Oct 07;23(10):e31374 [FREE Full text] [doi: 10.2196/31374] [Medline: 34516389]
2. National contact point: Health boards. NHS Inform. 2022 Jan 28. URL: https://www.nhsinform.scot/care-support-and-rights/health-rights/european-cross-border-healthcare/national-contact-point-health-boards [accessed 2021-11-04]
3. Rural Scotland: key facts 2018. Government of Scotland. 2018 Oct 08. URL: https://www.gov.scot/publications/rural-scotland-key-facts-2018/pages/2/ [accessed 2021-11-04]
4. Digital health and care strategy. Government of Scotland. 2021 Oct 27. URL: https://www.gov.scot/publications/scotlands-digital-health-care-strategy/ [accessed 2021-11-04]

Edited by T Leung; submitted 04.11.21; this is a non–peer-reviewed article;accepted 14.01.22; published 03.02.22.

Please cite as:
Weatherburn C
Digital Inclusion as a Foundation for Health Equity. Comment on "Expanding Video Consultation Services at Pace and Scale in Scotland During the COVID-19 Pandemic: National Mixed Methods Case Study"
J Med Internet Res 2022;24(2):e34702
URL: https://www.jmir.org/2022/2/e34702
doi:10.2196/34702
PMID:35038303

Letter to the Editor

Authors' Reply: Understanding the Impact of Social Media Information and Misinformation Producers on Health Information Seeking. Comment on "Health Information Seeking Behaviors on Social Media During the COVID-19 Pandemic Among American Social Networking Site Users: Survey Study"

Stephen Neely[1], PhD; Christina Eldredge[2], MD, PhD; Ronald Sanders[3], DPA

[1]School of Public Affairs, University of South Florida, Tampa, FL, United States

[2]School of Information, University of South Florida, Tampa, FL, United States

[3]Florida Center for Cybersecurity, School of Information, University of South Florida, Tampa, FL, United States

Corresponding Author:
Stephen Neely, PhD
School of Public Affairs
University of South Florida
4202 E Fowler Ave
SOC 107
Tampa, FL, 33620
United States
Phone: 1 8139748423
Email: srneely@usf.edu

Related Articles:

Comment on: https://www.jmir.org/2022/2/e31415

Comment on: https://www.jmir.org/2021/6/e29802

(J Med Internet Res 2022;24(2):e31569) doi:10.2196/31569

KEYWORDS

social media; internet; communication; public health; COVID-19; usage; United States; information seeking; web-based health information; online health information; survey; mistrust; vaccination; misinformation

We appreciate Boudreau and colleagues' [1] thoughtful consideration of our recent survey study [2], which examined American people's use of social networking sites (SNS) to learn and stay informed about the COVID-19 pandemic. As they point out, we surveyed a representative sample of American adults (N=1003) and found that most SNS users had not fact-checked COVID-19–related information with a medical professional, and those who had opted to follow credible, scientific sources on social media were significantly more likely to undergo vaccination [2]. In reply, Boudreau and colleagues noted that our study—and others like it—has focused primarily on consumers rather than the producers and publishers of medical content on social media [1]. They propose that researchers should shift their focus "from the consumers to the producers of this information," and, in particular, they emphasize the possibility of developing tools to assess and classify health-related posts on social media in order to help consumers distinguish medically valid guidance from potential misinformation.

We understand and affirm the underlying spirit of Boudreau et al's [1] recommendation, and building on that, we would endorse an "all of the above" approach to the study of social media moving forward. A comprehensive research agenda—drawing on a diverse range of perspectives and methodological techniques—will be needed in order to understand and keep pace with social media's growing and evolving role in health information seeking. This includes greater attention to issues of content and publisher credibility, as the authors suggest, though it should be noted that social media often obscures the distinction between publishers and consumers [3]. It also means that health professionals will need to gain awareness of and interpret emerging techniques in data mining, natural language processing, and network analysis. These are essential to identifying influential network nodes and understanding how

health information spreads in complex social networks. For reference, we conducted a similar analysis during the 2015-2016 Zika virus outbreak [4].

However, in pursuing a comprehensive research agenda around social media, it is critical that researchers not lose sight of the consumer perspective. We agree that promoting and affirming accuracy "at the source" is critical, but so too is understanding which sources of health information consumers encounter, trust, and rely on. Unfortunately, recent studies have noted declining trust in science among many Americans, including the Institution of Medicine [5,6]. This is especially salient in the case of politicized public health emergencies, such as the COVID-19 pandemic. Add to this the politicization and fragmentation of social media platforms themselves, and we find ourselves immersed in an information environment where even quality markers are often interpreted as political statements. While health professionals are not to blame for these trends, it is nonetheless important that they be aware of and responsive to them. This means that it is critical for research and scholars to stay focused on understanding consumer-level preferences, behaviors, and outcomes while also working to improve health messaging at its source.

Acknowledgments

The original research referred to in this letter was sponsored by the Florida Center for Cybersecurity at the University of South Florida.

Conflicts of Interest

None declared.

References

1. Boudreau HS, Singh N, Boyd CJ. Understanding the Impact of Social Media Information and Misinformation Producers on Health Information Seeking. Comment on "Health Information Seeking Behaviors on Social Media During the COVID-19 Pandemic Among American Social Networking Site Users: Survey Study". J Med Internet Res 2022 Jan 25:2022 [FREE Full text] [doi: 10.2196/31415] [Medline: 35076408]
2. Neely S, Eldredge C, Sanders R. Health Information Seeking Behaviors on Social Media During the COVID-19 Pandemic Among American Social Networking Site Users: Survey Study. J Med Internet Res 2021 Jun 11;23(6):e29802 [FREE Full text] [doi: 10.2196/29802] [Medline: 34043526]
3. Hughes AL, Tapia AH. Social media in crisis: When professional responders meet digital volunteers. J Homel Secur 2020;12(3):679-706. [doi: 10.1515/jhsem-2014-0080]
4. Hagen L, Keller T, Neely S, DePaula N, Robert-Cooperman C. Crisis Communications in the Age of Social Media. Soc Sci Comput Rev 2017 Aug 21;36(5):523-541. [doi: 10.1177/0894439317721985]
5. Baron RJ, Berinsky AJ. Mistrust in Science - A Threat to the Patient-Physician Relationship. N Engl J Med 2019 Jul 11;381(2):182-185. [doi: 10.1056/NEJMms1813043] [Medline: 31291524]
6. Blendon RJ, Benson JM, Hero JO. Public trust in physicians--U.S. medicine in international perspective. N Engl J Med 2014 Oct 23;371(17):1570-1572. [doi: 10.1056/NEJMp1407373] [Medline: 25337746]

Abbreviations

SNS: social networking sites

Edited by T Leung; submitted 25.06.21; this is a non–peer-reviewed article;accepted 25.01.22; published 04.02.22.

Please cite as:
Neely S, Eldredge C, Sanders R
Authors' Reply: Understanding the Impact of Social Media Information and Misinformation Producers on Health Information Seeking. Comment on "Health Information Seeking Behaviors on Social Media During the COVID-19 Pandemic Among American Social Networking Site Users: Survey Study"
J Med Internet Res 2022;24(2):e31569
URL: https://www.jmir.org/2022/2/e31569
doi:10.2196/31569
PMID:35119376

Letter to the Editor

Getting a Vaccine, Jab, or Vax Is More Than a Regular Expression. Comment on "COVID-19 Vaccine-Related Discussion on Twitter: Topic Modeling and Sentiment Analysis"

Jack Alexander Cummins[1]

Manchester Essex Regional High School, Manchester, MA, United States

Corresponding Author:
Jack Alexander Cummins
Manchester Essex Regional High School
36 Lincoln Street
Manchester, MA, 01944
United States
Phone: 1 9788101169
Email: 2jackcummins@gmail.com

Related Article:

Comment on: https://www.jmir.org/2021/6/e24435

(J Med Internet Res 2022;24(2):e31978) doi:10.2196/31978

KEYWORDS

COVID-19; vaccine; vaccination; Twitter; infodemiology; infoveillance; topic; sentiment; opinion; discussion; communication; social media; perception; concern; emotion; natural language processing

Lyu et al [1] used natural language processing techniques to analyze the topics and sentiment of Twitter conversations related to the COVID-19 vaccine using the Twitter chatter data set created by Georgia State University's Panacea Lab. Tweets that contained any of the following keywords "vaccination," "vaccinations," "vaccine," "vaccines," "immunization," "vaccinate," and "vaccinated" were selected for further analysis (eg, topic modeling and sentiment analysis).

I would suggest that the study might be enhanced by including additional keyword searches for vaccination synonyms identified by a method such as the Continuous Bag of Words (CBOW) word2vec model [2]. While the study focused on formal language such as "vaccination" and "vaccine," other colloquial terms are commonly used on Twitter and other social media platforms to describe the vaccination process.

I identified synonyms for "vaccination" commonly used on Twitter by using the *gensim* implementation of the CBOW word2vec model [3]. This model predicts synonyms and related words by creating vector representations of words. Words with similar vector representations are more likely to be synonyms than words with dissimilar vector representations. I trained the CBOW word2vec model on 503,862 tweets containing the keyword "covid" or "corona" from June 24-27, 2021, collected through the *rtweet* package [4]. The keyword pattern search results included tweets using words related to COVID-19 such as "covid-19" or "coronavirus."

Out of the 503,862 COVID-19–related tweets downloaded with *rtweet*, 94,768 contained at least one of the words searched for in the study by Lyu et al [1]. In addition, a total of 22,587 tweets used the terms "shot," "shots," "jab," "jabs," "jabbed," "vax," or "vaxxed." The words "shot" or "shots" were used in 9017 tweets. The words "jab," "jabs," or "jabbed" were used in 7021 tweets. The words "vax" or "vaxxed" were used in 4081 tweets. Out of the 22,587 tweets that contained these alternative terms, 15,855 (70.2%) were tweeted by users who self-disclosed their location on their user profile. Using the Nominatim application programming interface, it was possible to identify geocoded location, including country, for 13,101 of the 15,855 user-disclosed locations [5]. Of these 13,101 geocoded tweets, 3111 were from the United Kingdom, of which 2261 used "jab," "jabbed," or "jabs." Among the geocoded tweets, 4910 were from the United States; of these, 2704 included "shot" or "shots" and 1130 used "vax" and "vaxxed."

I would propose that researchers performing keyword searches on social media chatter consider using the CBOW word2vec model to enhance their studies by expanding the number of comments they capture and to reduce geographic or population bias that may occur from the preselection of terminology. The CBOW word2vec model can help capture more completely the full range of word choices used by social media users.

Conflicts of Interest

None declared.

Editorial Notice

The corresponding author of *"COVID-19 Vaccine–Related Discussion on Twitter: Topic Modeling and Sentiment Analysis"* declined to respond to this letter.

References

1. Lyu JC, Han EL, Luli GK. COVID-19 Vaccine-Related Discussion on Twitter: Topic Modeling and Sentiment Analysis. J Med Internet Res 2021 Jun 29;23(6):e24435 [FREE Full text] [doi: 10.2196/24435] [Medline: 34115608]
2. Mikolov T, Chen K, Corrado G, Dean J. Efficient Estimation of Word Representations in Vector Space. arXiv. Preprint posted online Jan 16, 2013. [FREE Full text]
3. Řehůřek R, Sojka P. Software Framework for Topic Modelling with Large Corpora. 2010 Presented at: Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks; May 22; Valletta, Malta URL: https://radimrehurek.com/lrec2010_final.pdf [doi: 10.13140/2.1.2393.1847]
4. Kearney M. rtweet: Collecting and analyzing Twitter data. JOSS 2019 Oct;4(42):1829. [doi: 10.21105/joss.01829]
5. Nominatim. OpenStreetMap Wiki. URL: https://wiki.openstreetmap.org/wiki/Nominatim [accessed 2021-07-12]

Abbreviations

CBOW: Continuous Bag of Words

Edited by T Leung; submitted 12.07.21; this is a non–peer-reviewed article;accepted 01.02.22; published 23.02.22.

Please cite as:
Cummins JA
Getting a Vaccine, Jab, or Vax Is More Than a Regular Expression. Comment on "COVID-19 Vaccine-Related Discussion on Twitter: Topic Modeling and Sentiment Analysis"
J Med Internet Res 2022;24(2):e31978
URL: https://www.jmir.org/2022/2/e31978
doi:10.2196/31978
PMID:35195531

Viewpoint

Control Centre for Intensive Care as a Tool for Effective Coordination, Real-Time Monitoring, and Strategic Planning During the COVID-19 Pandemic

Martin Komenda[1,2,3], PhD; Vladimír Černý[4,5], MD, PhD; Petr Šnajdárek[6], MSc; Matěj Karolyi[1,2], MSc; Miloš Hejný[1,2], BSc; Petr Panoška[1,2], BSc; Jiří Jarkovský[1,2], PhD; Jakub Gregor[1,2], PhD; Vojtěch Bulhart[1,2]; Lenka Šnajdrová[1,2], PhD; Ondřej Májek[1,2], PhD; Tomáš Vymazal[7], MD, PhD; Jan Blatný[4,8], MD, PhD; Ladislav Dušek[1,2], PhD

[1]Institute of Biostatistics and Analyses, Faculty of Medicine, Masaryk University, Brno, Czech Republic

[2]Institute of Health Information and Statistics of the Czech Republic, Prague, Czech Republic

[3]Department of Simulation Medicine, Faculty of Medicine, Masaryk University, Brno, Czech Republic

[4]Ministry of Health of the Czech Republic, Prague, Czech Republic

[5]Clinic of Anaesthesiology, Perioperative and Intensive Medicine, Masaryk Hospital in Ústí nad Labem, Ústí nad Labem, Czech Republic

[6]General Staff, Czech Armed Forces, Prague, Czech Republic

[7]Clinic of Anaesthesiology, Resuscitation and Intensive Medicine, University Hospital in Motol, Second Faculty of Medicine of Charles University, Prague, Czech Republic

[8]Department of Paediatric Haematology and Biochemistry, University Hospital Brno, Brno, Czech Republic

Corresponding Author:
Martin Komenda, PhD
Institute of Biostatistics and Analyses
Faculty of Medicine
Masaryk University
Kamenice 126/3
Brno, 62500
Czech Republic
Phone: 420 549 49 4469
Email: komenda@iba.muni.cz

Abstract

In the Czech Republic, the strategic data-based and organizational support for individual regions and for providers of acute care at the nationwide level is coordinated by the Ministry of Health. At the beginning of the COVID-19 pandemic, the country needed to very quickly implement a system for the monitoring, reporting, and overall management of hospital capacities. The aim of this viewpoint is to describe the purpose and basic functions of a web-based application named “Control Centre for Intensive Care,” which was developed and made available to meet the needs of systematic online technical support for the management of intensive inpatient care across the Czech Republic during the first wave of the pandemic in spring 2020. Two tools of key importance are described in the context of national methodology: one module for regular online updates and overall monitoring of currently free capacities of intensive care in real time, and a second module for online entering and overall record-keeping of requirements on medications for COVID-19 patients. A total of 134 intensive care providers and 927 users from hospitals across all 14 regions of the Czech Republic were registered in the central Control Centre for Intensive Care database as of March 31, 2021. This web-based application enabled continuous monitoring and decision-making during the mass surge of critical care from autumn 2020 to spring 2021. The Control Center for Intensive Care has become an indispensable part of a set of online tools that are employed on a regular basis for crisis management at the time of the COVID-19 pandemic.

(J Med Internet Res 2022;24(2):e33149) doi:10.2196/33149

KEYWORDS

COVID-19; coronavirus; intensive care; inpatient care; online control center; prescription; open data; ICU; monitoring; strategy; development; app; function; Czech Republic; inpatient; crisis management

Introduction

The COVID-19 pandemic has placed an urgent burden on health care systems, communication among their components, and health care management. Although the majority of patients have a mild or even asymptomatic course of the disease, approximately 8% to 15% of patients require hospitalization during the infection or afterward [1]. Approximately 21% to 36% of patients hospitalized due to COVID-19 are admitted to the intensive care unit (ICU) [2-4], and the average mortality of those admitted to the ICU is approximately 30% [5,6].

In the majority of developed countries, ICU beds are often close to their full capacity even under normal circumstances, let alone during the COVID-19 pandemic. Modeling studies suggest that a pandemic of severe influenza or another similar disease would require ICU and mechanic ventilation capacity that is significantly greater than what is available and that many patients who would require a ventilator might not have access to one [7,8].

By definition, epidemics and disasters will result in many patients arriving in a continuous stream, with shortages of necessary technology, beds, oxygen support, ventilators, medications, as well as trained health care personnel. The high likelihood of such conditions emphasizes the importance of having a framework that is suitably constructed to allow users to anticipate and adapt to these inevitable complexities and challenges [8]. Beds, equipment, and medications should be monitored ideally in “real time” and electronically, with knowledge of both their absolute numbers and whereabouts immediately available to system leaders [9]. For example, ventilatory support is an absolute necessity for the survival of critically ill patients and may be the single most important therapy that dictates the outcomes. Moreover, ventilators are highly demanding for staff qualification; they are therefore likely to be the limiting factor in any hospital’s ability to accommodate a large surge of mass critical care [10]. The COVID-19 pandemic has generated a sense of urgency that will allow the adoption of innovation without the logistical barriers and path dependencies that we have become accustomed to. Web-based solutions can fill a critical gap for the allocation of health care resources [11].

The Czech National Control Centre for Acute Inpatient Care (CNCC-AIC) was established under the auspices of the Ministry of Health of the Czech Republic in May 2020 to provide strategic data-based and organizational support for individual regions and for providers of acute care at the nationwide level. Its main purpose is to monitor hospital capacities, to provide analytical reporting on these capacities, and to enable the overall management of health care facilities run by all inpatient care providers in the Czech Republic.

The operative online tools of key importance employed by the CNCC-AIC involve the Control Centre for Intensive Care (CC-IC) [12], an online database updated on a daily basis, which provides an overview of capacities of inpatient care (classified into categories of acute, long-term, and supportive care) and a dedicated “Clinic” module, which is part of the Information System of Infectious Diseases (a database of people with laboratory-confirmed COVID-19).

In accordance with the defined methodology, the CC-IC was brought into operation in April 2020 to enable the continuous monitoring as well as daily reporting of available data on the occupancy rate of inpatient beds. The aim of this viewpoint paper is to present the methodology of managing capacities in Czech hospitals during the ongoing COVID-19 pandemic. A newly developed and robust technical solution, namely the CC-IC, is also introduced. The methodological background of the design and implementation of this unique practical operative tool, which is built on the long-term experience of experts across the Czech health care system, and experience with its practical use are presented.

Structure of the CC-IC

Overview

The CC-IC is an entirely new platform that has been developed since the start of the COVID-19 pandemic. This has been achieved in mutual cooperation among the following groups of involved stakeholders and experts to assess recent progress of the epidemic and recommend measures for crisis situations caused by the COVID-19 epidemic in the Czech Republic: (1) a team of regional coordinators of intensive care; (2) the Integrated Central Management Team, which includes representatives from the Ministry of Health, Czech Armed Forces, National Institute of Public Health, National Agency for Communication and Information Technologies, and Institute of Health Information and Statistics of the Czech Republic; and (3) a team of developers from a joint workplace of the Institute of Health Information and Statistics of the Czech Republic and the Institute of Biostatistics and Analyses at the Faculty of Medicine of Masaryk University.

The CC-IC provides two tools of key importance to offer effective support for the CNCC-AIC: (1) a module for regular online updates and overall monitoring of currently free capacities in hospitals (health care technology/medical devices, beds, staff) in real time, and (2) a module for online entering and overall record-keeping of requirements on medications for COVID-19 patients. No patient personal data or medical records are stored in the CC-IC application; thus, no anonymization or pseudonymization algorithms are used.

The CC-IC also specifically monitors the reprofiled capacity (ie, beds that had previously, under normal circumstances, in a given health care facility) been intended for the provision of care of another type or another specialty.

The CC-IC is a web application that has been continuously developed and run by the Institute of Health Information and Statistics of the Czech Republic, which cooperates closely with the Ministry of Health of the Czech Republic. Since the very beginning, the development has been based on requirements and needs for the coordination of intensive care, as required by health care facilities in the Czech Republic (Figure 1).

Use of the application (available in the Czech language only) relies on several basic principles: (1) user authentication and

authorization, (2) definition of user roles and rights to edit and read, (3) each new user registration is approved by the CNCC-AIC, and (4) private and free email addresses are not allowed. CC-IC users include representatives of individual health care facilities (management, coordinators of intensive care, physicians, and nurses), emergency medical service workers, pharmacy workers, and members of the Integrated Central Management Team.

The use case diagram in Figure 2 provides an overview of the entire CC-IC system, including key functions assigned to users in specific roles. The main objective of this diagram is to depict the various actors as well as interactions that are available to these actors.

Use cases provide insight into the basic structure of CC-IC functional requirements. Based on modeling of individual use cases, which are associated with the primary actors, we can see individual parts of the system so that we can decompose and divide the whole system into separate submodules. Moreover, the CC-IC provides an authenticated application programming interface (API) for secure sharing of selected data sets, which consists of access to a list of requirements on medications and an export of currently free capacities in hospitals, including their history.

Figure 1. Schematic diagram of the Control Centre for Intensive Care (CC-IC).

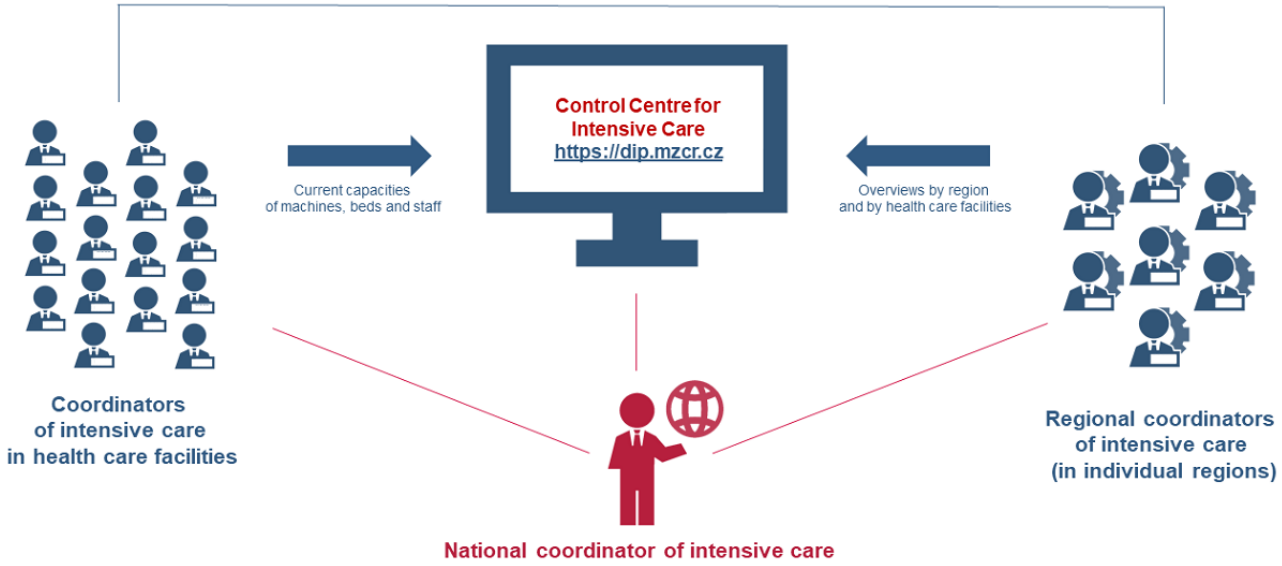
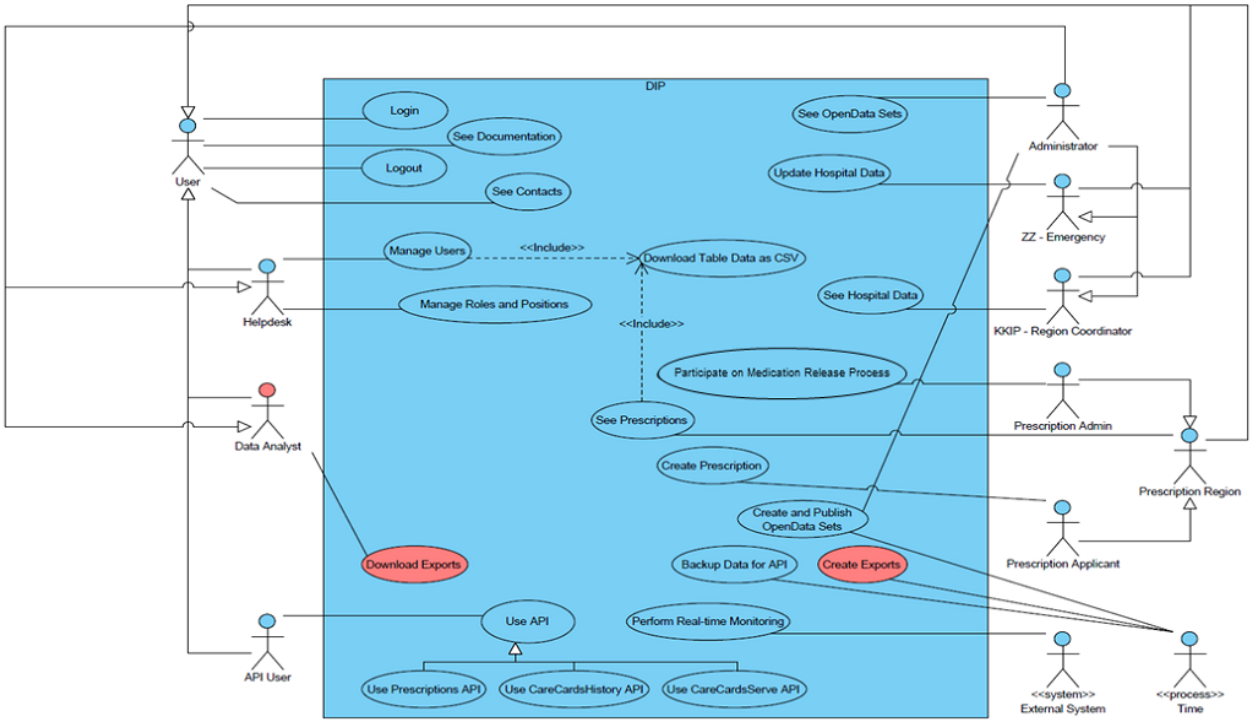


Figure 2. Use case diagram for the Control Centre for Intensive Care system. API: application programming interface.



Regular Monitoring of Health Care Providers and the ICU Network

On April 9, 2020, the Ministry of Health of the Czech Republic adopted an exceptional measure, ordering all acute care providers to use the CC-IC to immediately report any changes in capacities of intensive care related to COVID-19 patients. Since the launch of the CC-IC, the availability of free capacity of machines, beds, and staff has been monitored on a regular basis (Textbox 1). Hospitals have been requested to report their intensive care capacities on a daily basis, usually at the end of the day. The system allows for continuous updates at any time; more frequent entering of the latest data would nevertheless bring an unnecessary burden to hospital staff at the time of a significant shortage of capacities in terms of personnel, time, and equipment.

Based on data in the CC-IC, it is possible to coordinate the provision of inpatient care centrally, on a nationwide level. In each individual region, hospitalizations of patients who need admission to an intensive care facility (ICU and/or emergency department) are managed and coordinated by the so-called regional coordinator of intensive care, who is a physician specialized in anesthesiology and intensive care medicine. These physicians cooperate closely with regional emergency medical services and with representatives of regional authorities. Apart from that, each health care facility has prepared emergency beds (ie, inpatient beds that could be used to provide care for COVID-19 patients if all existing capacities are occupied).

The system reacts to immediate needs by individual regions, depending on the current occupancy rate of machines and beds, combined with the number of COVID-19 patients. If an increase in the number of new patients is detected within the CC-IC, it is necessary to prepare health care facilities for a crisis situation. This involves limiting capacities for patient admissions (ie, only admit patients who require intensive and emergency care, and suspend the provision of nonemergency care); discharging as many patients as possible to home care; plan an increased bed supply in health care facilities for an increase in COVID-19 patients; and provide a sufficient number of competent health care professionals.

The module for regular online updates and overall monitoring of the currently free capacities of intensive care in real time is predominantly used by coordinators of intensive care in individual health care facilities. A total of 134 intensive care providers from all 14 regions of the Czech Republic are registered in the central CC-IC database. As of March 31, 2021, a total of 927 users were registered who used this module to update data on intensive care capacities in predefined intervals (ie, at least once a day during critical periods of the COVID-19 pandemic), who made 86,565 sessions in total.

A user-friendly and simple interface (Figure 3) is available to users, who can therefore update occupancy rates of available capacities (in terms of machines, beds, and staff) very easily. The entire application is fully responsive and therefore entirely compatible with all types of devices, namely smartphones, tablets, laptops, and desktops. Free capacities for COVID-19–positive versus COVID-19–negative patients are distinguished only for beds. Entering a free bed into the system means that this bed is fully functional (ie, all required equipment and corresponding health care professionals are available). Monitoring of staff and equipment in and out of service has turned out to be ineffective and unsustainable on a long-term basis; this function was therefore removed from the system upon the decision of the Central Integrated Management Team. Across the entire module, each individual card has one of three background colors to provide a quick overview of free capacities: white, <30% to 100%> free capacity; yellow, <10% to 30% free capacity; and red, <0% to 10% of free capacity.

Textbox 1. Overview of monitored information on capacities of health care facilities providing inpatient care in the Czech Republic.

Health care technology/medical devices

- Extracorporeal membrane oxygenation
- Mechanical ventilation in accident and emergency (A&E) departments and in intensive care units (ICUs) for adults
- Continuous renal replacement therapy
- Intermittent renal replacement therapy
- Transport ventilators
- Anesthesia machine with mechanical ventilator

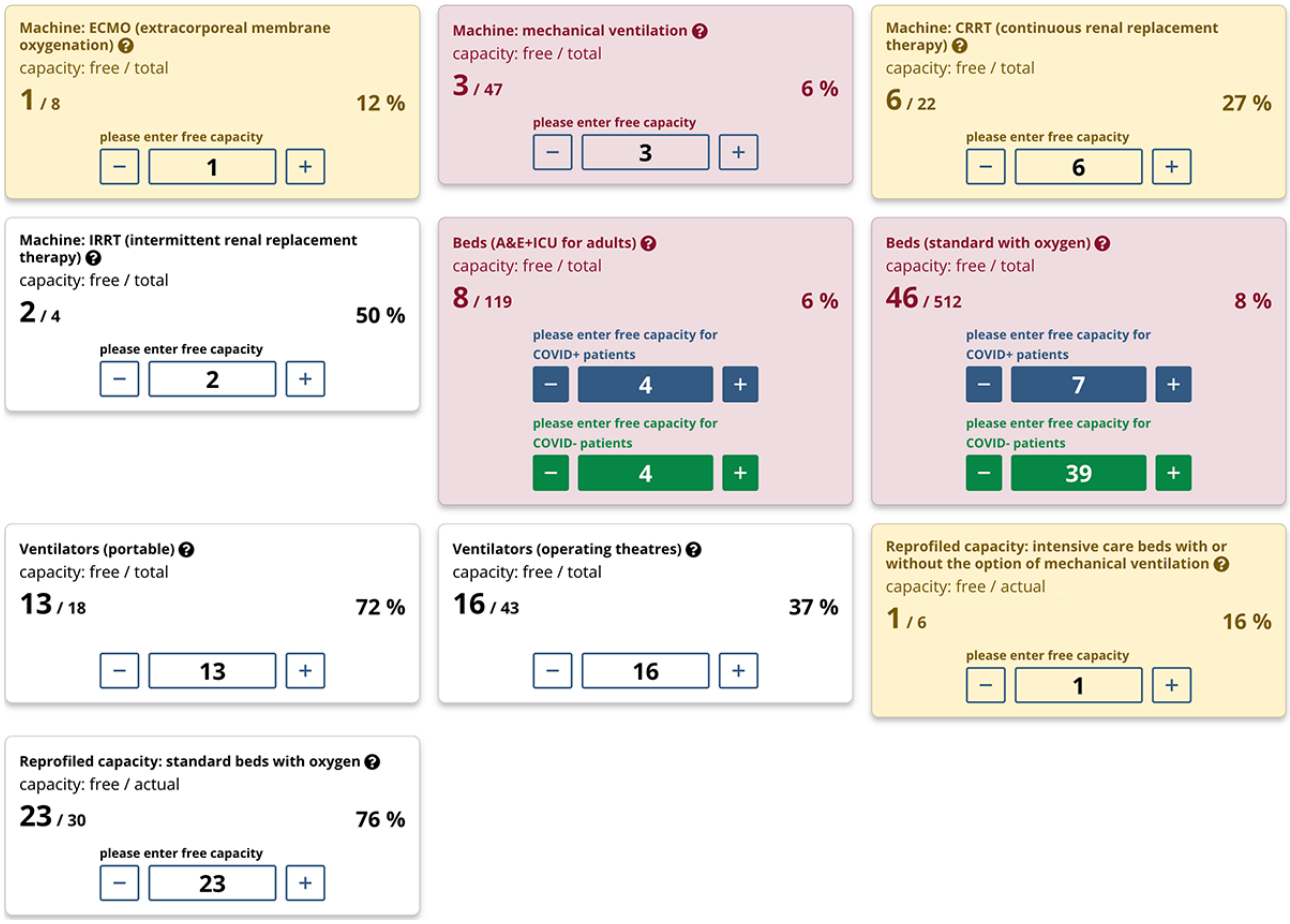
Hospital beds

- A&E and ICU beds for adults; a distinction is made between beds for COVID-19–positive and COVID-19–negative patients
- Standard beds with oxygen; a distinction is made between beds for COVID-19–positive and COVID-19–negative patients
- Reprofiled capacity: beds with or without the option of ventilatory support devices

Staff

- Physicians (A&E and ICU for adults)
- Nurses (A&E and ICU for adults)

Figure 3. Interface for updates of free capacities. Overview of intensive care capacities for the capital, Prague, as of December 30, 2020. A&E: accident and emergency; ICU: intensive care unit.



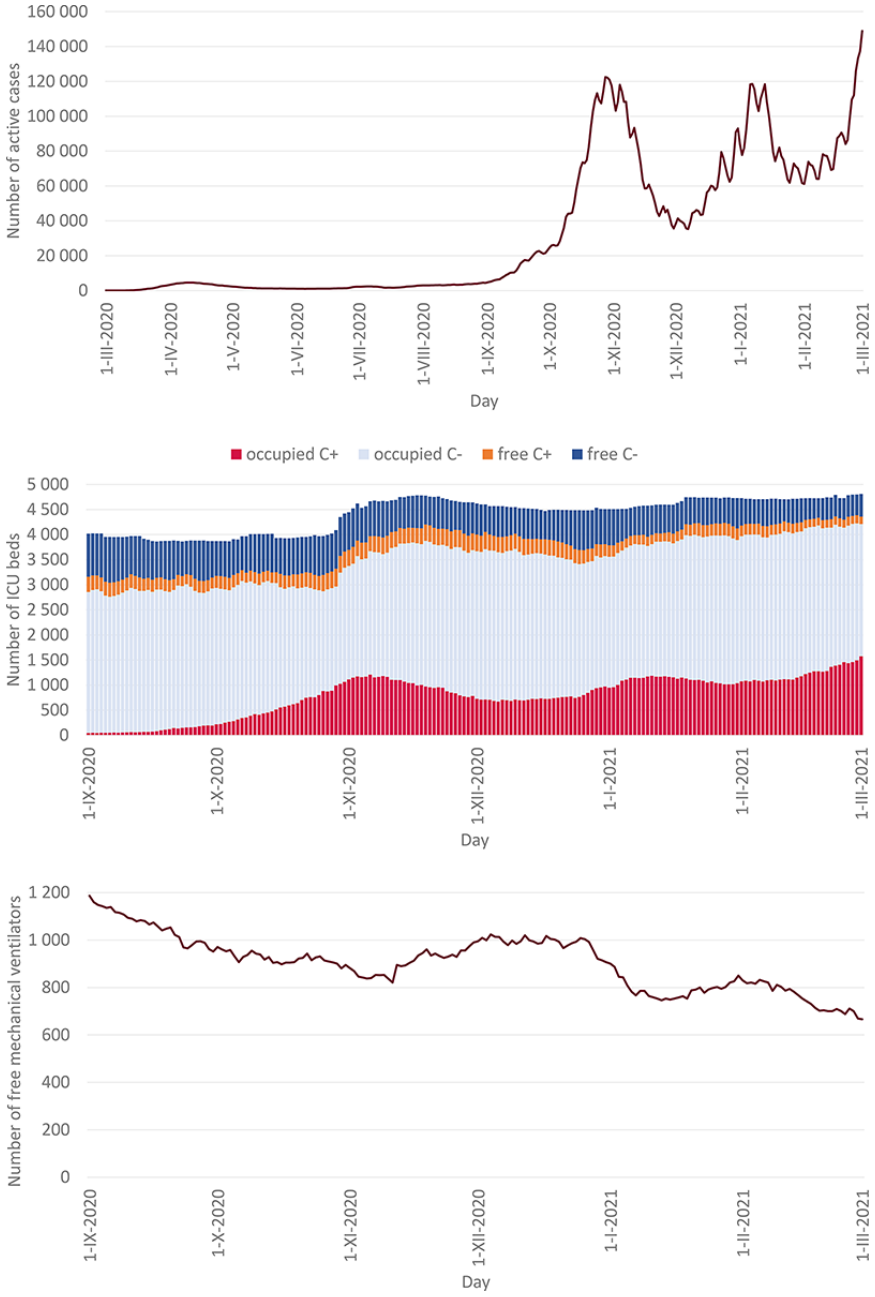
Two different reports are available to provide a quick overview of the latest data. These reports are primarily used by regional coordinators of intensive care, members of the operational team of the CNCC-AIC, members of management of individual health care facilities, and top officials of the Ministry of Health of the Czech Republic. The first report is a summary data table providing a clearly arranged view of all health care facilities, including real-time absolute and relative occupancy numbers as regard to available capacities. The second is an aggregated visualization in the form of context cards, which also show real-time absolute and relative occupancy numbers in regard to available capacities, but users can further choose among the nationwide view, views for individual regions, or those for individual health care facilities.

The example screenshot from the CC-IC application in Figure 3 also demonstrates an aggregated visualization of key parameters monitored in real time for the capital, Prague. Four cards have a yellow background because the available capacity of corresponding parameters was reported to be below 30% at the time.

Owing to the timely and strict measures at the national level, the first COVID-19 wave in the spring of 2020 did not bring significant pressure to hospital-based care, with a maximum daily prevalence of around 4500 active cases and less than 100 ICU beds being occupied with COVID-19 patients among the Czech population of over 10 million people. The situation changed in the autumn of 2020; the number of active cases peaked twice to approximately 120,000, first in early November 2020 and then in early January 2021 (Figure 4, top). The number of patients requiring intensive care at the ICU fluctuated at around 1200 in these periods. Requirements for intensive care increased to an even more critical level in the second half of February 2021, with 1574 ICU beds occupied by COVID-19 patients on March 1, 2021 (Figure 4, middle).

The increasing numbers of hospitalized patients with COVID-19 during October 2020, as steadily monitored by the CC-IC, required increased capacities of standard beds and ICU beds. The total numbers of both types of beds were increased by repurposing the existing beds and adding new beds, according to the capabilities of individual hospitals. The availability of mechanical ventilators as an essential equipment of intensive care followed the overall trends: it decreased from nearly 1200 machines on September 1, 2020, to less than 700 at the end of February 2021 (Figure 4, bottom). If a temporary depletion of ICU capacities occurred in some regions, patients were transferred to hospitals in other regions in which free capacities were available according to the CC-IC report. Over 100 patients were transferred between hospitals during autumn 2020 and winter 2021, usually from small local hospitals to large hospitals in regional capitals or in Prague. All of these decisions were taken by the national and regional intensive care coordinators in cooperation with the Central Integrated Management Team.

Figure 4. Time trends of COVID-19 cases and related intensive care, from top to bottom: number of active cases (prevalence), occupied and free intensive care unit (ICU) beds for COVID-19 patients and other patients, and availability of mechanical ventilators.



Prescription of COVID-19 Medications

With an increasing number of hospital admissions of COVID-19 patients who need medication to manage the disease, it was necessary to secure the systematic record-keeping and distribution of these medications to individual health care facilities. For this reason, a dedicated module was developed within the CC-IC, making it possible to request any of three therapies against COVID-19: remdesivir, convalescent plasma, and favipiravir. The tool has been intentionally designed in a rather flexible manner so that it can be easily adapted and employed after the end of the pandemic, thus providing a generally applicable ordering system for any medication.

The CC-IC enables physicians from any involved health care facility to request a certain medication for a specific patient. The module has three basic user roles: (1) applicant for medication (physician), which has been assigned to an approved group of registered users from individual health care facilities who are thus entitled to enter new requests from their respective health care facility; (2) regional coordinator of intensive care, who is entitled to enter new requests from health care facilities from across the entire region and can also see an overview of requests not only from the region but from the Czech Republic as a whole; (3) virtual indication group, with members entitled to approve requests for medications from the respective region and they can also see an overview of requests from across the Czech Republic; and (4) contact persons from pharmacies linked to specific health care facilities, who are entitled to supply the approved medications.

Mutual cooperation among representatives of the Ministry of Health of the Czech Republic, regional coordinators of intensive care, members of the virtual indication group, and

representatives of pharmacies led to the development of a structured form that is employed to send a request for a specific medication. The form consists of three parts: (1) health care facility, containing information on the physician and hospital requesting the medication; (2) specification of therapy, containing information on the requested medication and the date of supply; and (3) patient description, which is an anonymized description of the patient's condition and their risk factors (without mentioning any personal data in conformity with personal data protection according to General Data Protection Regulation).

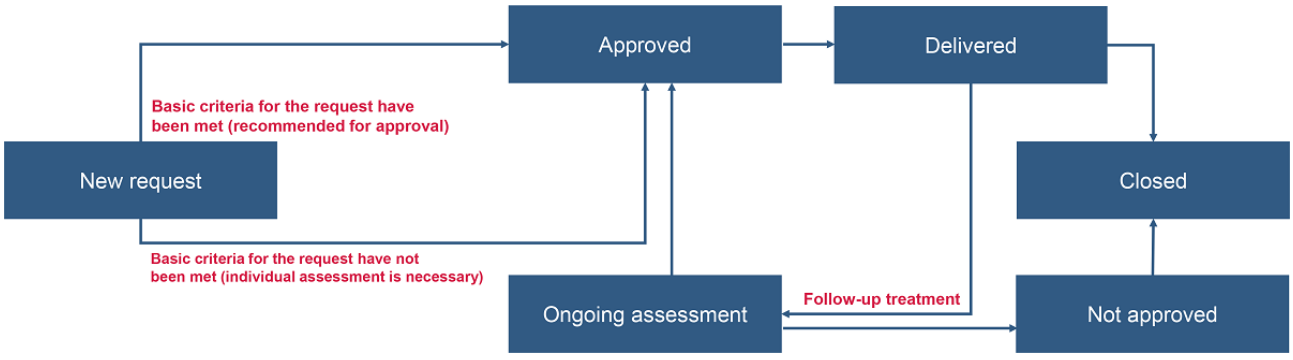
After the form is filled out and sent, validation of individual items is performed within the CC-IC, the request is stored in the central database, and a notification is sent to all users involved (physician, member of the virtual indication group in the respective region, pharmacy). The medication must be explicitly approved before being supplied by a respective pharmacy that secures supplies for a given health care facility. In the final step, the request is closed for the purpose of archiving and subsequent statement of charges for the respective health insurance company. In exceptional cases, request for follow-up treatment can occur after the administration of remdesivir; such requests can also be processed by the CC-IC. The life cycle of a request for a medication within the CC-IC is shown in Figure 5.

All health care facilities providing acute care began to report, by means of the CC-IC, the current use of these capacities on April 9, 2020, when the first version of the system was launched. The first version only contained one module, which enabled regular online updates and overall monitoring of currently free capacities of intensive care in real time. On June 10, 2020, the second module was launched for the online entering and overall record-keeping of requirements on medications for COVID-19 patients.

A major added value of the entire CC-IC system is the fact that it was designed to meet the current needs of acute medicine in response to the unexpected burden on inpatient care capacity caused by the COVID-19 pandemic. The data that the CC-IC allows to collect provide a key support for data-oriented decision-making processes within ICU occupancy crisis management. Thanks to the CNCC-AIC and a proper methodological setup, it is possible to effectively coordinate, for example, elective care, which can have a negative effect on patients' health when being moved or postponed. The following paragraphs briefly describe the functions of both of the above-mentioned modules as well as analytical reports and open data sets that have been provided as an outcome of the CC-IC database.

As of December 30, 2020, a total of 4759 requests for medication were registered in the CC-IC, out of which 4666 (98.05%) were requests for remdesivir. The remaining 93 (1.95%) requests were for convalescent plasma. On that date, a total of 496 users entitled to work with requests for medications were registered in the system; these users included physicians in individual health care facilities, regional coordinators of intensive care, pharmacy workers, and members of the virtual indication group. Between the first wave and the second wave of the COVID-19 pandemic, there was a major change in the functioning of the module for the sending and record-keeping of requests for medications. A new version was implemented in September 2020, fully supporting the entire process of requests for medications (entering, validation of entered data, sending, approval/rejection, supply, conclusion), including email notifications. The form itself for sending a request contains a list of key information necessary for subsequent approval and the supply of a specific medicine (see Multimedia Appendix 1).

Figure 5. Description of the life cycle of a request for a medication in the Control Centre for Intensive Care (CC-IC).



Practical Use of Open Data

Data on free capacities are available as records in the National Catalogue of Open Data [13,14] in a regularly updated data set that describes two waves of the COVID-19 pandemic in the Czech Republic.

The data set contains overviews of machine equipment (extracorporeal membrane oxygenation, mechanical ventilation, continuous renal replacement therapy, intermittent hemodialysis, ventilators—portable and those in operating theaters) and occupancy rate of inpatient beds (only beds in the emergency department and ICU for adults and standard beds with oxygen in the entire hospital). Data are regularly updated for beds suitable for the provision of the requested type of care (ie, necessary staff as well as corresponding material and technical equipment must also be available). The definition of the necessary number of staff for a given type of bed can be different among individual hospitals. These data have been available since September 1, 2020, when the obligation for hospitals to update free capacity was reset to a 1-day interval.

On top of that, for the purpose of the transfer of record-based (nonaggregated) data directly from the CC-IC to the Institute of Health Information and Statistics of the Czech Republic analytical team and to external authorized systems (run by the leadership of the Capital of Prague, for example), an API was developed to provide a machine-readable alternative to user access. For a secure transfer of data, each user must be authenticated and authorized, and the following conditions must be met: (1) each new user of the API must first register in the CC-IC system and must be explicitly approved for a given user role; (2) each request must contain the so-called API key (api_key), which is an integral part of the query (query string); (3) other parameters can also be part of the query string parameter, which are listed at specific endpoints; and (4) for each user, there is a defined list of allowed endpoints, IP addresses, and parameter values. If the user's request does not pass through the authorization process, the user will obtain the response "401 Unauthorized."

Data from the CC-IC, along with information from other registries (eg, the Information System of Infectious Diseases), are regularly processed for the purpose of daily reporting, which is provided to the team of the CNCC-AIC. Among dozens of analytical outputs (as of December 23, 2020), one of them monitors—on a daily basis—the situation in the Czech Republic from the points of view of currently hospitalized patients, new admissions to hospitals, new discharges from hospitals, as well as the utilization of health care facilities in individual regions, expressed by their free capacity [15].

Discussion

The Czech Republic was one of the most affected countries during the COVID-19 pandemic waves in autumn 2020 and winter 2021 in terms of incidence, hospitalizations, and deaths. The CC-IC was one of the key tools developed to manage the capacities of standard and ICU beds in the entire country, and to ensure that appropriate care was provided not only to patients with COVID-19 but also to other patients.

The Czech Republic has by far the highest density of ICU beds (43.2 beds per 100,000 population) and one of the longest average hospital stays (9.5 days) among the Organization for Economic Cooperation and Development countries [16]. There is also a high level of centralization of hospital-based care, because all large hospitals are under the control of either the Ministry of Health or regional authorities. These factors enabled a very quick response and management of intensive care, including the compulsory reporting of current capacities into the central system.

The spread of COVID-19 has usually not been uniform, and there have been different ICU demands within the affected countries and their regions. Central reporting systems for the monitoring of ICU capacities were therefore established in some countries [17]. The absence of such a platform, often combined with a fragmented health care system and small ICU facilities, may slow down the response and hold back the provision of appropriate care to critically ill patients, both with and without COVID-19 [18].

The systems for ICU capacities monitoring employed in various countries are usually operated under the auspices of national authorities (ie, government, ministries, and armed forces). They have been designed to provide near real-time data for crisis management and decision-making [19-21]. Short- and long-term estimates of ICU bed occupancy in the upcoming days and weeks using statistical models are another important output [21,22].

The CC-IC platform has been designed and developed in a tailor-made manner, responding to needs of the leadership of the Ministry of Health of the Czech Republic and crisis teams that are actively involved in the management of the COVID-19 epidemic in the Czech Republic. Based on the need to quickly and effectively resolve various crisis situations related to the capacity of ICUs, the CC-IC system was incorporated into the Czech legislation. Experience with the COVID-19 pandemic has clearly shown the requirement for long-term monitoring of ICU bed occupancy. This is the main reason why the obligation to update the occupancy on a daily basis for all health care facilities providing inpatient care in the Czech Republic is now enshrined in the Czech legislation. As part of international cooperation, the possibility of transfer of this online tool to other countries is being discussed. With respect to the general concept of the CC-IC design, the transferability of this tool to other countries should be trouble-free provided that the basic methodology for capacity monitoring in health care facilities is adhered to. It is of key importance that new requirements are collected from health care professionals themselves, even in the future. Intensivists should be part of strategic planning committees before, during, and after pandemics to coordinate ICU responses with hospital and regional efforts for triage, clinical care, and infection control [7]. Regular evaluation includes the analysis of user behavior and further optimization of the whole platform, so that it can continue to function as the primary support tool for all involved stakeholders.

Conclusions

The web platform CC-IC has been developed as a reaction to the urgent need of strategic data-based and organizational support for individual regions and providers of acute care on a nationwide level. The platform is intended for online real-time reporting of changes of free capacity of hospitals, real-time reporting of occupancy rate and availability of beds, a clearly arranged reporting of health care facilities, and the basis for quick decision-making of crisis management during the COVID-19 pandemic. Delegated representatives of health care facilities (coordinators of intensive care) are tasked with reporting the remaining free capacity of machines and beds that are available for patients hospitalized with COVID-19. Information on newly reported current capacity is immediately available to the Integrated Central Management Team, which continuously monitors the situation across the Czech Republic. Online entering and nationwide record-keeping of requests for medications intended for COVID-19 patients are processed by a standalone module. This system can be accessed by users from individual health care facilities providing acute care; the users typically involve several physicians and management representatives from the same health care facility, so that they can stand in for each other and share the latest information.

Development of the CC-IC is far from complete; the system has been further improved, partly based on requirements by members of the team of the CNCC-AIC and partly in response to suggestions by users themselves.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Key information necessary for subsequent approval of a specific medication.
[DOCX File , 15 KB - jmir_v24i2e33149_app1.docx]

References

1. Stawicki S, Jeanmonod R, Miller A, Paladino L, Gaieski D, Yaffee A, et al. The 2019-2020 novel coronavirus (Severe Acute Respiratory Syndrome Coronavirus 2) pandemic: a joint American College of Academic International Medicine-World Academic Council of Emergency Medicine Multidisciplinary COVID-19 Working Group Consensus Paper. J Glob Infect Dis 2020;12(2):47-93 [FREE Full text] [doi: 10.4103/jgid.jgid_86_20] [Medline: 32773996]
2. Boëlle PY, Delory T, Maynadier X, Janssen C, Piarroux R, Pichenot M, et al. Trajectories of hospitalization in COVID-19 patients: an observational study in France. J Clin Med 2020 Sep 29;9(10):3148 [FREE Full text] [doi: 10.3390/jcm9103148] [Medline: 33003375]
3. Gottlieb M, Sansom S, Frankenberger C, Ward E, Hota B. Clinical course and factors associated with hospitalization and critical illness among COVID-19 patients in Chicago, Illinois. Acad Emerg Med 2020 Oct;27(10):963-973. [doi: 10.1111/acem.14104] [Medline: 32762106]
4. Nachtigall I, Lenga P, Jóźwiak K, Thürmann P, Meier-Hellmann A, Kuhlen R, et al. Clinical course and factors associated with outcomes among 1904 patients hospitalized with COVID-19 in Germany: an observational study. Clin Microbiol Infect 2020 Dec;26(12):1663-1669 [FREE Full text] [doi: 10.1016/j.cmi.2020.08.011] [Medline: 32822883]
5. Abate SM, Ahmed Ali S, Mantfardo B, Basu B. Rate of intensive care unit admission and outcomes among patients with coronavirus: a systematic review and meta-analysis. PLoS One 2020 Jul 10;15(7):e0235653 [FREE Full text] [doi: 10.1371/journal.pone.0235653] [Medline: 32649661]
6. Serafim RB, Póvoa P, Souza-Dantas V, Kalil AC, Salluh JI. Clinical course and outcomes of critically ill patients with COVID-19 infection: a systematic review. Clin Microbiol Infect 2021 Jan;27(1):47-54 [FREE Full text] [doi: 10.1016/j.cmi.2020.10.017] [Medline: 33190794]
7. Kain T, Fowler R. Preparing intensive care for the next pandemic influenza. Crit Care 2019 Oct 30;23(1):337 [FREE Full text] [doi: 10.1186/s13054-019-2616-1] [Medline: 31665057]
8. Daugherty Biddison EL, Faden R, Gwon H, Mareiniss DP, Regenberg AC, Schoch-Spana M, et al. Too many patients…a framework to guide statewide allocation of scarce mechanical ventilation during disasters. Chest 2019 Apr;155(4):848-854. [doi: 10.1016/j.chest.2018.09.025] [Medline: 30316913]
9. Dichter J, Kanter R, Dries D, Luyckx V, Lim ML, Wilgis J, Task Force for Mass Critical Care. System-level planning, coordination, and communication: care of the critically ill and injured during pandemics and disasters: CHEST consensus statement. Chest 2014 Oct;146(4 Suppl):e87S-e102S [FREE Full text] [doi: 10.1378/chest.14-0738] [Medline: 25144713]
10. Einav S, Hick J, Hanfling D, Erstad BL, Toner ES, Branson RD, Task Force for Mass Critical Care. Surge capacity logistics: care of the critically ill and injured during pandemics and disasters: CHEST consensus statement. Chest 2014 Oct;146(4 Suppl):e17S-e43S. [doi: 10.1378/chest.14-0734] [Medline: 25144407]
11. Krausz M, Westenberg J, Vigo D, Spence R, Ramsey D. Emergency response to COVID-19 in Canada: platform development and implementation for eHealth in crisis management. JMIR Public Health Surveill 2020 May 15;6(2):e18995 [FREE Full text] [doi: 10.2196/18995] [Medline: 32401218]
12. COVID 19: online intensive care dispatching. Control Centre for Intensive Care. URL: https://dip.mzcr.cz/dip-login [accessed 2022-02-11]
13. Klímek J. DCAT-AP representation of Czech National Open Data Catalog and its impact. J Web Semant 2019 Mar;55:69-85. [doi: 10.1016/j.websem.2018.11.001]
14. National Catalogue of Open Data. URL: https://opendata.mzcr.cz/ [accessed 2022-02-11]
15. Komenda M, Bulhart V, Karolyi M, Jarkovský J, Mužík J, Májek O, et al. Complex reporting of the COVID-19 epidemic in the Czech Republic: use of an interactive web-based app in practice. J Med Internet Res 2020 May 27;22(5):e19367 [FREE Full text] [doi: 10.2196/19367] [Medline: 32412422]
16. Health at a Glance 2021: OECD Indicators. Paris: OECD Publishing; 2021.
17. Winkelmann J, Webb E, Williams G, Hernández-Quevedo C, Maier CB, Panteli D. European countries' responses in ensuring sufficient physical infrastructure and workforce capacity during the first COVID-19 wave. Health Policy 2021 Jul 09:online ahead of print. [doi: 10.1016/j.healthpol.2021.06.015] [Medline: 34311982]

18. Yamamoto T, Ozaki M, Kasugai D, Burnham G. Assessment of critical care surge capacity during the COVID-19 pandemic in Japan. Health Secur 2021;19(5):479-487. [doi: 10.1089/hs.2020.0227] [Medline: 34346775]
19. Pilcher D, Coatsworth N, Rosenow M, McClure J. A national system for monitoring intensive care unit demand and capacity: the Critical Health Resources Information System (CHRIS). Med J Aust 2021 Apr;214(7):297-298 [FREE Full text] [doi: 10.5694/mja2.50988] [Medline: 33774832]
20. Raposo G, Castellanos J, Hernandez H. Near real-time telemetry of nation-wide ICU's capacity strain in Argentina. Stud Health Technol Inform 2021 May 27;281:820-821. [doi: 10.3233/SHTI210291] [Medline: 34042787]
21. Zhao C, Tepekule B, Criscuolo NG, Wendel Garcia PD, Hilty MP, RISC-ICU Consortium Investigators In Switzerland, et al. icumonitoring.ch: a platform for short-term forecasting of intensive care unit occupancy during the COVID-19 epidemic in Switzerland. Swiss Med Wkly 2020 May 04;150:w20277. [doi: 10.4414/smw.2020.20277] [Medline: 32374886]
22. de Salles Neto LL, Martins CB, Chaves A, Roma de Oliveira Konstantyner TC, Yanasse HH, de Campos CBL, et al. Forecast UTI: application for predicting intensive care unit beds in the context of the COVID-19 pandemic. Epidemiol Serv Saude 2020;29(4):e2020391 [FREE Full text] [doi: 10.5123/S1679-49742020000400023] [Medline: 32997068]

Abbreviations

API: application programming interface
CC-IC: Control Centre for Intensive Care
CNCC-AIC: Czech National Control Centre for Acute Inpatient Care
ICU: intensive care unit

Edited by C Basch; submitted 25.08.21; peer-reviewed by P Štourač, M Henry, C Guinemer; comments to author 20.11.21; revised version received 10.12.21; accepted 20.12.21; published 16.02.22.

Please cite as:
Komenda M, Černý V, Šnajdárek P, Karolyi M, Hejný M, Panoška P, Jarkovský J, Gregor J, Bulhart V, Šnajdrová L, Májek O, Vymazal T, Blatný J, Dušek L
Control Centre for Intensive Care as a Tool for Effective Coordination, Real-Time Monitoring, and Strategic Planning During the COVID-19 Pandemic
J Med Internet Res 2022;24(2):e33149
URL: https://www.jmir.org/2022/2/e33149
doi:10.2196/33149
PMID:34995207

Original Paper

The Impact of COVID-19 Confinement on Cognition and Mental Health and Technology Use Among Socially Vulnerable Older People: Retrospective Cohort Study

Elena Dura-Perez[1,2*], NP; Jessica Marian Goodman-Casanova[1*], PMHCNS; Amanda Vega-Nuñez[1*], CP; Gloria Guerrero-Pertiñez[1], PsyD; Esperanza Varela-Moreno[1], PsyM; Maite Garolera[3], PhD; Maria Quintana[3], PhD; Antonio I Cuesta-Vargas[4], PhD; Pilar Barnestein-Fonseca[1], PhD; Carlos Gómez Sánchez-Lafuente[1*], Psych; Fermin Mayoral-Cleries[1*], PhD; Jose Guzman-Parra[1*], PhD

[1]Department of Mental Health, Regional University Hospital of Málaga, Biomedical Research Institute of Malaga (IBIMA), Málaga, Spain

[2]Faculty of Psychology, University of Málaga, Málaga, Spain

[3]Brain, Cognition and Behavior: Clinical Research, Consorci Sanitari de Terrassa, Terrasa, Spain

[4]Department of Physiotherapy, University of Málaga, Biomedical Research Institute of Malaga, Málaga, Spain

[*]these authors contributed equally

Corresponding Author:
Jessica Marian Goodman-Casanova, PMHCNS
Department of Mental Health
Regional University Hospital of Málaga
Biomedical Research Institute of Malaga (IBIMA)
Plaza del Hospital, s/n
Málaga, 29009
Spain
Phone: 34 660901966
Email: jmariangoodman@gmail.com

Abstract

Background: COVID-19 forced the implementation of restrictive measures in Spain, such as lockdown, home confinement, social distancing, and isolation. It is necessary to study whether limited access to basic services and decreased family and social support could have deleterious effects on cognition, quality of life, and mental health in vulnerable older people.

Objective: This study aims to explore the impact of the COVID-19 outbreak on cognition in older adults with mild cognitive impairment or dementia as the main outcome and the quality of life, perceived health status, and depression as secondary outcomes and to analyze the association of living alone and a change in living arrangements with those outcomes and other variables related with the use of technology and health services. Likewise, this study aims to analyze the association of high and low technophilia with those variables, to explore the access and use of health care and social support services, and, finally, to explore the informative-, cognitive-, entertainment-, and socialization-related uses of information and communications technologies (ICTs) during the COVID-19 outbreak.

Methods: This cohort study was conducted in Málaga (Spain). In total, 151 participants with mild cognitive impairment or mild dementia, from the SMART4MD (n=75, 49.7%) and TV-AssistDem (n=76, 50.3%) randomized clinical trials, were interviewed by telephone between May 11 and June 26, 2020. All participants had undergone 1-3 assessments (in 6-month intervals) on cognition, quality of life, and mood prior to the COVID-19 breakout.

Results: The outbreak did not significantly impact the cognition, quality of life, and mood of our study population when making comparisons with baseline assessments prior to the outbreak. Perceived stress was reported as moderate during the outbreak. After correction for multiple comparisons, living alone, a change in living arrangements, and technophilia were not associated with negative mental health outcomes. However, being alone was nominally associated with self-perceived fear and depression, and higher technophilia with better quality of life, less boredom, perceived stress and depression, and also less calmness. Overall, health care and social support service access and utilization were high. The most used ICTs during the COVID-19 outbreak were the television for informative, cognitive, and entertainment-related uses and the smartphone for socialization.

Conclusions: Our findings show that the first months of the outbreak did not significantly impact the cognition, quality of life, perceived health status, and depression of our study population when making comparisons with baseline assessments prior to the outbreak. Living alone and low technophilia require further research to establish whether they are risk factors of mental health problems during lockdowns in vulnerable populations. Moreover, although ICTs have proven to be useful for informative-, cognitive-, entertainment-, and socialization-related uses during the pandemic, more evidence is needed to support these interventions.

Trial Registration: ClinicalTrials.gov NCT04385797; https://clinicaltrials.gov/ct2/show/NCT04385797

International Registered Report Identifier (IRRID): RR2-10.2196/26431

(J Med Internet Res 2022;24(2):e30598) doi:10.2196/30598

KEYWORDS

COVID-19; cognition; quality of life; social isolation; mental health; social support; technology; physical distancing; leisure activities; nursing

Introduction

COVID-19 was declared a worldwide pandemic by the World Health Organization on March 11, 2020 [1]. To avoid the serious collapse of health systems in response to the rising number of cases and deaths, European countries, as in other continents, decided to implement different measures to control the pandemic.

In Spain, the government decided to declare a national state of alarm, implementing restrictive measures from March 15 until June 21, 2020. The measures included lockdown, home confinement, social distancing, and isolation (activities were limited to basic needs, such as buying food or medication, attending health care centers and financial institutions); closure of schools and nonessential activities; ban of all internal travels except for essential ones; and border closure [2]. These measures also led to a change in health care access: Only critical attention was guaranteed, patient care changed from on-site interviews to telephonic attention, visits with medical specialists were suspended, and there was a lack of monitoring of chronic pathologies.

The elderly population is 1 of the groups most socially vulnerable to this disease. Age alone is by far the most significant factor for death due to COVID-19 [3]. Although COVID-19 infects people of all ages, the risk of becoming seriously ill increases in adults aged over 40 years, and especially in those aged over 60 years. In Spain, 68% of all hospitalizations due to coronavirus and 95% of all deaths correspond to the population over the age of 60 years, with a notable increase after the age of 80 years [4].

Recent data suggest that in addition to old age and medical comorbidities (eg, hypertension, diabetes, obesity), dementia is associated with an increased risk of having severe COVID‐19 and related mortality [5-8]. In addition, restrictive measures, such as confinement, may pose a risk for people with mild cognitive impairment (MCI) or mild dementia (MD). COVID-19 confinement has resulted in an increase in known risk factors for dementia, such as inactivity [9], limited access to basic services [10], isolation [11], and decreased family and social support [12]. These factors could have deleterious effects on cognition, quality of life, and overall health [13]. Therefore, the pandemic has not only a health impact on people with MCI/MD but also a social impact.

Loneliness and social isolation often coexist and are all too common in older adults. Loneliness refers to the subjective state of feeling alone, separated, or apart from others. Social isolation, in contrast, is defined as the objective physical separation from other people, such as living alone, in which one has few social relationships or there is a low frequency of interaction with others [14].

Considering the latter definition, we can understand that the COVID-19 pandemic has increased the social isolation of older adults as restrictive measures have enforced staying at home, distancing, and shutting down all nonessential activities. This has meant that people have been forced to minimize their social interactions to avoid the spread of the virus, leaving those who live alone completely isolated. Social isolation has been identified as a health risk factor as it reduces well-being and is associated with higher prevalence of depression [15] and cognitive impairment [16]. In older adults, it has a greater impact due to decreased social resources, functional and mobility limitations, death of family members, and changes in family structures [17].

During quarantine, factors such as boredom and a lack of activities play an important role. They can contribute to depression [18] and have an impact on the quality of life and functional dependence [19]. Mental activity, in contrast, may improve cognitive function and reduce overall dementia risk [20].

In the “information age,” information and communications technologies (ICTs) have emerged for combating loneliness and social isolation [21]. Although the age-related digital divide and health-related conditions (cognitive, visual, motor, etc) may compromise the use of technologies in the elderly, the extensive home penetration of ICTs has facilitated remote, home-based interventions. These interventions reduce the risk of viral exposure and prevent health-related negative outcomes of social isolation through health care delivery, cognitive stimulation, social connection, information sharing, and leisure entertainment [22].

The aims of this study were (1) to explore the impact of the COVID-19 outbreak on cognition in community-dwelling older

adults with MCI/MD as the main outcome and the quality of life, perceived health status, and depression as secondary outcomes; (2) to analyze the differences between individuals living alone and living with others regarding mental health, and other variables related with the use of technology and health services during the COVID-19 outbreak and, likewise, to explore the effect of a change in living arrangements on cognition, quality of life, perceived health status, and depression; (3) to analyze the differences between individuals with high and low technophilia regarding mental health and other variables related with use of technology and health services during the COVID-19 outbreak; (4) to explore the access and use of health care and social support services during the COVID-19 outbreak; and finally (5) to explore the informative-, cognitive-, entertainment-, and socialization-related uses of ICTs during the COVID-19 outbreak.

Methods

Study Design

This cohort study was conducted in the Spanish region of Málaga (Andalucía) and approved by the North-East Malaga Ethics Committee (1078-N-20). Interviews were telephone-administered to guarantee the safest means to communicate during the COVID-19 pandemic. Researchers contacted participants by telephone, explained the study in detail, answered any questions that arose, and obtained consent from those willing to participate in the study [23].

Ethics Approval and Consent to Participate

The study was approved by the North-East Malaga Ethics Committee (1078-N-20). Participants provided written consent before taking part.

Trial Registration

This study was registered in ClinicalTrials.gov (NCT04385797).

Setting

Participants were identified from the Support, Monitoring and Reminder Technology for Mild Dementia (SMART4MD; NCT03325699) [24] and TV-Based Assistive Integrated Service to Support European Adults Living with Dementia (TV-AssistDem; NCT03653234) [25] randomized clinical trials (RCTs), which aimed to assess the effects of ICTs to support MCI/MD using a tablet-based health application and a television-based assistive integrated service, respectively. In both RCTs, a broad definition of MCI, a subjective memory deterioration sustained over time, was considered. All participants had undergone 1-3 previous assessments (in 6-month intervals) in the RCTs on cognition, quality of life, and depression prior to the COVID-19 breakout.

Participants

Researchers from the Biomedical Research Institute of Malaga contacted 210 potential respondents from the SMART4MD (n=111, 52.9%) and TV-AssistDem (n=99, 47.1%) RCTs by telephone. In total, 151 participants, SMART4MD (n=75, 49.7%) and TV-AssistDem (n=76, 50.3%), agreed to participate. However, for 8 (5.3%) of them, it was not possible to assess the main variable (cognition) and the secondaries variables (quality of life, perceived health status, and depression), because their abilities to answer the questionnaires were compromised during the time of assessment.

Participants were eligible for inclusion when the following criteria applied: participating in the SMART4MD and TV-AssistDem RCTs and agreeing to participate by giving consent. Eligibility criteria of the aforementioned RCTs were age>55 years or >60 years, perception of memory problems for at least 6 months, score of 20-28 or 23-27 points in the Mini-Mental State Exam (MMSE), independently living, having an informal caregiver, and taking care of their medical prescription. Patients with a score above 11 on the Geriatric Depression Scale (GDS), a terminal illness, or specific cognitive or physical conditions that would reduce their ability to use a tablet or a television were excluded.

Interview Process

Participants were contacted by telephone by 5 health care professionals (2 neuropsychologists, 1 clinical psychologist, 1 psychologist, and 1 psychiatric and mental health clinical nurse specialist). Researchers had previously established relationships with participants during both RCTs. Quantitative and qualitative strategies were used to create an unstandardized ad hoc telephone-based survey in order to gather as much information as soon as possible. The exceptional situation did not allow us to test the instrument prior to its implementation by phone. To minimize the interference of this situation in the results, validated phone versions tests were used.

The survey was a useful tool for guiding the interviewers and gathering information simultaneously in a homogenous way. A model of the questionnaire used is attached in Annex 1 in Multimedia Appendix 1.

Researchers interviewed the participants between May 11 and June 26, 2020. The variables of sociodemographic data (age, sex, and living arrangements), health perception-management (change in living arrangements due to lockdown, presence of COVID-19 symptoms, frequency of access to COVID-19 information), sleep-rest patterns, types of ICTs (smartphone, tablet, television, laptop), and their uses (informative, cognitive, entertainment, and socialization) were collected from the participants unless their abilities to answer such a long interview were compromised, in which case the caregivers were interviewed on their behalf. The questionnaires that evaluated the main variables (cognition, quality of life, depression, perceived stress, and technophilia) were answered by the participants.

The mean time from the start of the lockdown and home confinement measures to the interview was 70.36 days (SD 12.40, range 52-102).

Instruments Used Before and After the Lockdown

Cognition

The primary outcome variable was cognition. During the assessment prior to the COVID-19 outbreak (T0), the MMSE [26] was used to assess the cognitive function of the participants with MCI/MD. We decided to use as eligibility criteria a broad

spectrum, because although the common cut-off score for cognitive impairment is 24, it has been shown that an MMSE cut-off score of 28 provides high sensitivity and specificity for detecting MD in a well-educated population with self-reported memory complaints [27].

During the COVID-19 outbreak (T1), the validated telephone version of the MMSE had to be used to maintain health and safety measures. This phone version has a maximum score of 22 because it cannot cover all sections [28]. For example, on the spatial orientation section, researchers were not able to check on which floor the patient was. Motor skills or some language skills could not be measured either. In the telephone version, the subject is asked only to repeat a phrase and name 1 item. The items of the original version, such as naming a second word, asking to follow a 3-stage command, reading and obeying a sentence, writing a sentence, or copying an intersecting pentagon, could not be measured.

Although the full version of the MMSE was used in the T0 assessment, for data analysis, the scoring was based on the 22 items of the phone version.

Quality of Life and Perceived Health Status

The health-related quality of life (HRQoL) of the participants was measured in both assessments using the total score of the Quality of Life-Alzheimer's Disease Scale (QoL-AD) [29]. The QoL-AD is a 13-item measure, in which responses are 4-point multiple-choice options (1=poor, 2=fair, 3=good, 4=excellent). It includes questions related to the interpersonal, environmental, functional, physical, and psychological status of a person with dementia, and thus, it is a global measure for the quality of life. Scale scores range from 13 to 52, with higher scores indicating a greater quality of life. In cases where patients had compromised cognitive function, informal caregivers completed the QoL-AD in parallel and on behalf of the people with MCI/MD.

The European Quality of Life 5 Dimensions 3 Levels (EuroQoL-5D-3L) [30] was also administered in both assessments. Currently, the EuroQoL-5D-3L is 1 of the most widely used generic preference-based measures in the world. It assesses an individual's HRQoL [31]. It has been shown to be valid in different patient groups and settings [32], including patients with cognitive impairment and dementia [33].

The EuroQoL-5D-3L consists of 5 questions along with a visual analog scale (VAS). The VAS records the patient's self-rated health on a vertical scale, where the endpoints are "the best health you can imagine" and "the worst health you can imagine." Due to the impossibility of the patients to see the VAS during the T1 assessment, they were asked to rate their health status. Only the VAS-perceived health status assessment was used for this study, combined with the Qol-AD.

Depression

The short form of the GDS was used during the T0 assessment [34]. It is a scale with 15 items and a range of scores, where 0-4 is considered normal, 5-8 indicates mild depression, 9-11 indicates moderate depression, and 12-15 indicates severe depression.

During the COVID-19 lockdown (T1), the telephone version of the GDS [35] was used. This version has high internal consistency and is highly correlated with the validated face-to-face administration of the scale, indicating that it is a valid instrument for screening depression among elderly people in special situations, such as the COVID-19 outbreak.

Instruments Used Only After the Lockdown

Technophilia

To measure older people's attitudes and enthusiasm toward technologies, the Instrument for Measuring Older People's Attitudes Toward Technology (TechPH) was used during the T1 assessment [36]. This questionnaire is designed to specifically assess technophilia in the older population and includes 6 items assessed on a 5-point Likert scale from 1 (fully disagree) to 5 (fully agree). The scale has 2 factors to define technophilia: technology enthusiasm and technology anxiety. It refers to a person's enthusiasm and positive feelings toward their technology use and the absence of fears and doubts about their ability to manage it.

Perceived Stress

The Perceived Stress Scale (PSS) [37] measures the degree to which situations in one's life are appraised as stressful. The scale has 14 questions regarding feelings and thoughts during the past month and are rated according to frequency (0=never, 1=almost never, 2=sometimes, 3=fairly often, 4=very often). PSS scores are obtained by reversing responses (eg, 0=4, 1=3, 2=2, 3=1, 4=0) to the 7 positively stated items (items 4, 5, 6, 7, 9, 10, and 13) and then summing across the entire 14-scale item. A higher score indicates a higher level of perceived stress. This questionnaire was used only during the T1 assessment. Studies report that the reduced version, PSS-10, has optimal psychometric properties in the general population and people exposed to confinement to assess perceived stress [38,39].

Other Variables

Other variables were sociodemographic data, including age, sex, and living arrangements; health perception-management (ie, change in living arrangements due to lockdown, presence of COVID-19 symptoms, frequency of access to COVID-19 information); coping-stress tolerance (ie, self-perceived mental health and well-being and mood); sleep-rest patterns (ie, self-perceived alterations in usual sleep patterns); and types of ICTs (smartphone, tablet, television, laptop) and their uses (informative, cognitive, entertainment, and socialization).

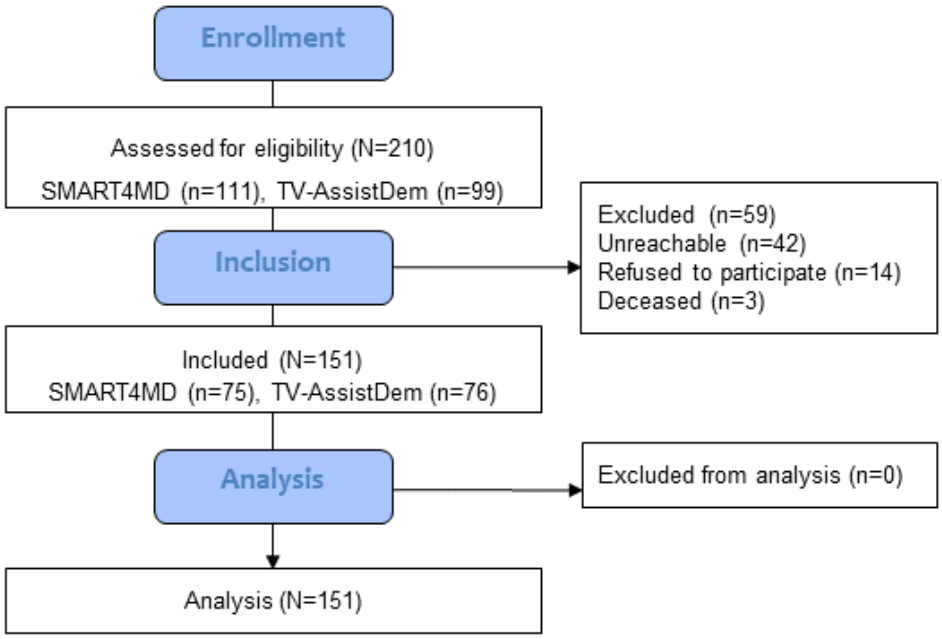
The survey data followed Gordon's Functional Health Patterns (Multimedia Appendix 1) [40]. To collect data on health perception, health management, and sleep-rest patterns, questions with numerically rated items were used. Open questions were included for the qualitative assessment of the patterns of coping-stress tolerance, activity-exercise, and role-relationship.

Data Analysis

The flow of participants is shown schematically with counts in a participant flow diagram (Figure 1). Statistics considered for presentation for continuous measures were the mean and SD, and if the criterion of normality was not met, the median and

the first and third quartiles. Categorical variables were summarized using counts and percentages.

Figure 1. Participant flow diagram. SMART4MD: Support, Monitoring and Reminder Technology for Mild Dementia; TV-AssistDem: TV-based ASSistive Integrated Service to supporT European adults living with mild DEMentia or mild cognitive impairment.



The change in means in the main outcome (cognition) and in the secondary outcomes (quality of life, perceived health status, and depression) were analyzed with respect to the last assessment of the RCTs (SMART4MD and TV-AssistDem) using the repeated measure *t* test or the nonparametric Wilcoxon test, if appropriate (considering significant values of α<.05). In the secondary outcomes, we applied Bonferroni correction for 3 comparisons (considering significant values of α<.017).

For the analysis of groups based on living arrangements (living alone vs living with others), the *t* test was used for continuous variables and the chi-square test for categorical variables. To evaluate the association of this variable with mental health outcomes, multivariate logistic regression for binary variables or linear regression for continuous variables was performed to adjust for confounders: age, sex, and technophilia. In addition, we explored the association between a change in living arrangements during the pandemic and cognition, depression, quality of life, and perceived health status with univariate and multivariate linear regression (adjusting for the following confounders: sex, age, and current living arrangements). For this comparison, we applied Bonferroni correction for 4 comparisons (considering significant values of α<.013).

For the analysis using groups based on the score in technophilia (based on the median of the TechPH index as the cut-off point), the *t* test or nonparametric Wilcoxon test was used for continuous variables and the chi-square test for categorical variables. To evaluate the association of this variable with mental health outcomes, multivariate logistic regression for binary variables or linear regression for continuous variables was performed to adjust for confounders: age, sex, and living arrangements.

To establish the Bonferroni correction for multiple comparisons (regarding living arrangements and technophilia groups), the number of independent tests was estimated with principal component analysis. Of the 37 variables analyzed, the first 35 components explained >99% (99.3%) of the variance. Thus, we considered the values significant with α<.0014.

To analyze the assumptions of all linear regression models, the Ramsey RESET linearity test, the Breusch-Pagan homoscedasticity test, and the Shapiro-Wilk normality test of the model residuals were used (see Annex 2 in Multimedia Appendix 1). When cognition, quality of life, health status, and depression were used as dependent variables, the assumption of the normality of the residuals of the model was not fulfilled, and transformation did not solve the problem, robust linear regression models were used (using the *robustbase* package and the *lmrob* function in R).

The R version 4.0.4 program was used for all statistical analysis [41].

Results

Participants

Of the 210 potential respondents (n=111 [52.9%] from SMART4MD and n=99 [47.1%] from TV-AssistDem), a total of 165 (78.6%) of 210 respondents was successfully reached, of which 151 (91.5%) agreed to participate (Figure 1). In addition, 150 (99.3%) participants completed the full interview without assistance, and 15 informal caregivers were interviewed on behalf of participants whose abilities to answer such a long interview were compromised. Of these 15 participants, 11 (73%) answered the main variables questionnaires and 4 (27%) were unable to do so. Given the complexity of MCI/MD, to not skew

the results of patients with a greater decline, all data were considered.

The mean time between the last assessment of the RCTs (T0) and the interview during the lockdown (T1) was 199.33 days (SD 52.43, range 67-395), and the mean duration of the calls was 50.14 minutes (SD 16.40).

Sociodemographics

The mean age of the sample was 74.31 years (SD 6.48), 97 (64.2%) of 151 participants were women, 36 (23.8%) lived alone, and 80 (53.3%) had high attraction to technology (high technophilia). The COVID-19 outbreak forced 22 (14.6%) of 151 participants to change their living arrangements (Table 1).

Table 1. Sample sociodemographic characteristics and differences between living alone and living with others, and high technophilia and low technophilia.

Characteristics	Total participants (N=151)	Living alone (n=36)	Living with others (n=115)	Statistics	*P* value	High technophilia (n=80)	Low technophilia (n=65)	Statistics	*P* value
Age (years), mean (SD)	74.31 (6.48)	76.31 (5.38)	73.69 (6.69)	t_{93}=–2.14	.03	73.69 (6.33)	74.74 (6.63)	t_{93}=0.97	.34
Sex, n (%)									
Male	54 (35.8)	4 (11.1)	50 (43.5)	χ^2_1=12.50	<.001	31 (63.3)	18 (36.7)	χ^2_1=1.96	.16
Female	97 (64.2)	32 (88.9)	65 (56.5)	χ^2_1=12.50	<.001	49 (51.0)	47 (49.0)	χ^2_1=1.96	.16
Change in living arrangements, n (%)									
Yes	22 (14.6)	6 (16.7)	16 (13.9)	χ^2_1=0.17	.68	11 (13.8)	10 (15.4)	χ^2_1=0.08	.78
No	129 (85.4)	30 (83.3)	99 (86.9)	χ^2_1=0.17	.68	69 (86.3)	55 (84.6)	χ^2_1=0.08	.78

Differences in Cognition, Quality of Life, Perceived Health Status, and Depression Prior to and During the COVID-19 Outbreak

Regarding the differences between the period before and during the outbreak, there were no differences in the main outcome: cognition. After correction for multiple comparisons, there were no statistically significant differences in the quality of life, perceived health status, or depression between the 2 periods (Table 2).

Table 2. Differences in cognition, quality of life, perceived health status, and depression prior to and during the COVID-19 outbreak.

Outcomes	Before the COVID-19 outbreak	During the COVID-19 outbreak	Statistics	*P* value
Main outcome				
Cognition (MMSE[a]), median (IQR)	19 (17-20)	19 (17-21)	Z=–0.798	.43
Secondary outcomes[b]				
QoL-AD[c], mean (SD)	35.97 (4.74)	36.25 (5.44)	t_{144}=–0.80	.43
Perceived health status (EuroQoL-5D-3L[d] thermometer), median (IQR)	70 (50-80)	70 (60-85)	Z=–1.94	.05
Depression (GDS[e]), median (IQR)	3 (1-5)	2 (1-4)	Z=–0.01	.99

[a]MMSE: Mini-Mental State Exam.

[b]Significant results with *P*<.02.

[c]QoL-AD: Quality of Life-Alzheimer's Disease Scale.

[d]EuroQoL-5D-3L: European Quality of Life 5 Dimensions 3 Levels.

[e]GDS: Geriatric Depression Scale.

Differences Between Individuals Living Alone and Living With Others, and With Change in Living Arrangements

Regarding social isolation (living alone and living with others), after Bonferroni correction, there was no significant association between the variables of the study (Table 3). Some factors reached nominally significance: self-perceived fear (being alone 36.1% vs 18.4%; χ^2=4.27; *P*=.04) and depression (being alone 3 vs 2; Z=–2.10; *P*=.04). In the multivariate models, depression did not reach nominally statistically significance (B=0.83; *P*=.06); see Table 3 (note: significant results with *P*<0.001). After Bonferroni correction, there were statistically significant differences regarding home visits, with those living alone (being alone 82.9% vs 51.3%; χ^2=10.97; *P*<.001) receiving more visits than those living with others. There was no significant

association regarding other variables (only a nominal significant association regarding the more frequent use of the newspaper in those living with others).

The change in living arrangements was not associated with cognition (unadjusted model: B=–0.21, *P*=.65; adjusted model: B=0.36, *P*=.44), quality of life (unadjusted model: B=–0.74, *P*=.40; adjusted model: B=1.05, *P*=.21), perceived health status (unadjusted model: B=3.34, *P*=.41; adjusted model: B=3.34, *P*=.41), or depression (unadjusted model: B=–0.15, *P*=.74; adjusted model: B=–0.25, *P*=.56). After correction for multiple comparisons, there was no association with less perceived stress (unadjusted model: B=–4.72, *P*=.01; adjusted model: B=–0.26, *P*=.02) but the results reached nominal significance. Annex 2 in Multimedia Appendix 1 shows more detailed information about linear regression models.

Table 3. Health perception-management, coping-stress tolerance, and sleep-rest functional health patterns during the COVID-19 outbreak and differences between living alone and with others.

Overall health status	Total participants (N=151)	Living alone (n=36)	Living with others (n=115)	Statistics	*P* value	Odds ratio (OR)/B[a]	*P* value
Health status (COVID-19), n (%)							
No symptoms	147 (97.4)	35 (97.2)	112 (97.4)	χ^2_4=4.13	.13	—[b]	—
Symptoms without test	3 (2.0)	0	3 (2.6)	χ^2_4=4.13	.13	—	—
Symptoms and positive test	1 (0.7)	1 (2.8)	0	χ^2_4=4.13	.13	—	—
Hospitalized	0	0	0	χ^2_4=4.13	.13	—	—
Intensive care unit (ICU) inpatient	0	0	0	χ^2_4=4.13	.13	—	—
Deceased	0	0	0	χ^2_4=4.13	.13	—	—
Self-perceived mental health and well-being, n (%)							
Well	108 (71.5)	23 (63.9)	85 (73.9)	W_1=1.27	.26	1.64	.27
Calm	64 (42.7)	14 (38.9)	50 (43.9)	W_1=1.20	.27	1.47	.37
Sad	49 (32.7)	15 (41.7)	34 (29.8)	W_1=0.05	.83	.91	.83
Worried	69 (46.0)	20 (55.6)	49 (43.0)	W_1=0.68	.41	.65	.31
Afraid	34 (22.7)	13 (36.1)	21 (18.4)	W_1=4.27	.04	.37	.04
Anxious	33 (22.0)	9 (25.0)	24 (21.1)	W_1=0.31	.58	.71	.49
Bored	28 (18.7)	9 (25.0)	19 (16.7)	W_1=1.84	.18	.46	.14
Self-perceived sleep quality, n (%)							
Maintained	117 (77.5)	29 (80.6)	88 (76.5)	W_1=0.12	.73	OR=1.07	.90
Altered	34 (22.5)	7 (19.3)	27 (23.5)	W_1=0.12	.73	OR=1.07	.90
Cognition (MMSE[c]), median (IQR)	19 (17-21)	19 (17-21)	19 (17-20.75)	Z=–0.57	.57	B[d]=0.30	.52
Quality of life (QoL-AD[e], mean (SD)	36.25 (5.44)	35.03 (4.39)	36.66 (5.70)	t_{143}=1.53	.13	B[d]=–1.88	.02
Perceived health status (EuroQoL-5D-3L[f] thermometer), median (IQR)	70 (60-85)	75 (60-100)	70 (60-80)	Z=–1.51	.13	B[d]=6.67	.12
Depression (GDS[g]), median (IQR)	2(1-4)	3 (2-5)	2 (1-4)	Z=–2.10	.04	B[d]=0.83	.06
Perceived stress (PSS[h]), mean (SD)	19.5 (8.64)	20.44 (7.96)	19.19 (8.87)	t_{149}=–0.75	.45	B[i]=0.08	.43

[a]Multivariate models (logistic or lineal) with living arrangements (living alone and living with others) as the independent variable and gender, age, and technophilia (high technophilia and low technophilia) as covariates. More information about linear regression models is shown in Annex 2 in Multimedia Appendix 1.

[b]Not applicable.

[c]MMSE: Mini-Mental State Exam.

[d]Robust linear regression.

[e]QoL-AD: Quality of Life-Alzheimer's Disease Scale.

[f]EuroQoL-5D-3L: European Quality of Life 5 Dimensions 3 Levels.

[g]GDS: Geriatric Depression Scale.

[h]PSS: Perceived Stress Scale.

[i]As the residuals of the model were not normal, we transformed the dependent variable in its logarithmic form.

Differences Between High- and Low-Technophilia Groups

After correction for multiple comparisons, there was no significant association between technophilia and the variables of the study. Only some variables reached nominal significant associations: self-perceived boredom (high technophilia 10.1% vs 27.7%; χ^2=7.44; *P*=.01), calmness (high technophilia 31.6% vs 52.3%; χ^2=6.30; *P*=.01), perceived stress (high technophilia 18.1 vs 21.23; *t*=2.19; *P*=.03), depression (high technophilia 2 vs 3; *Z*=2.16; *P*=.03), and quality of life (high technophilia 37.3

vs 35.3; t=2.24; P=.03). In the multivariate models, after controlling for possible confounders, the associations only maintained nominally statistical significance and the perceived health status reached nominal significance (B=6.44, P=.04); see Table 4 (note: significant results with P<.001).

Table 4. Health perception-management, coping-stress tolerance, and sleep-rest functional health patterns during the COVID-19 outbreak and differences between the high- and low-technophilia groups.

Overall health status	Total participants (N=151)	High technophilia (n=80)	Low technophilia (n=65)	Statistics	*P* value	Odds ratio (OR)/B[a]	*P* value
Health status (COVID-19), n (%)							
No symptoms	147 (97.4)	78 (97.5)	63 (96.9)	χ^2_2=1.39	.50	—[b]	—
Symptoms without test	3 (2.0)	2 (2.5)	1 (1.5)	χ^2_2=1.39	.50	—	—
Symptoms and positive test	1 (0.7)	0	1 (1.5)	χ^2_2=1.39	.50	—	—
Hospitalized	0	0	0	χ^2_2=1.39	.50	—	—
Intensive care unit (ICU) inpatient	0	0	0	χ^2_2=1.39	.50	—	—
Deceased	0	0	0	χ^2_2=1.39	.50	—	—
Self-perceived mental health and well-being, n (%)							
Well	108 (71.5)	54 (76.9)	50 (67.5)	χ^2_1=1.57	.21	1.64	.20
Calm	64 (42.7)	25 (31.6)	34 (52.3)	χ^2_1=6.30	.01	2.23	.02
Sad	49 (32.7)	22 (27.8)	25 (38.5)	χ^2_1=1.83	.18	1.41	.36
Worried	69 (46.0)	31 (39.2)	35 (53.8)	χ^2_1=3.06	.08	1.75	.11
Afraid	34 (22.7)	17 (21.5)	16 (24.6)	χ^2_1=0.194	.66	1.21	.65
Anxious	33 (22.0)	13 (16.5)	18 (27.7)	χ^2_1=2.67	.10	2.01	.10
Bored	28 (18.7)	8 (10.1)	18 (27.7)	χ^2_1=7.44	.01	3.69	.01
Self-perceived sleep quality, n (%)							
Maintained	117 (77.5)	66 (82.5)	47 (72.3)	χ^2_1=2.17	.14	1.92	.11
Altered	34 (22.5)	14 (17.5)	18 (27.7)	χ^2_1=2.17	.14	1.92	.11
Cognition (MMSE[c]), median (IQR)	19 (17-21)	19 (17-21)	18 (16.25-21)	Z=–1.13	.26	B[d]=0.30	.52
Quality of life (QoL-AD[e]), mean (SD)	36.25 (5.44)	37.33 (5.48)	35.33 (4.90)	t_{139}=2.24	.03	B[d]=1.64	.03
Perceived health status (EuroQoL-5D-3L[f] thermometer), median (IQR)	70 (60-85)	80 (60-90)	70 (60-80)	t_{142}=–1.65	.10	B[d]=6.44	.04
Depression (GDS[g]), median (IQR)	2(1-4)	2 (1-4)	3 (1-5)	Z=–2.16	.03	B[d]=–0.83	.03
Perceived stress (PSS[h]), mean (SD)	19.5 (8.64)	18.1 (8.77)	21.23 (8.21)	t_{141}=2.19	.03	B[i]=–0.19	.02

[a]Multivariate models (logistic or lineal) with technophilia (high and low) as the independent variable and gender, age, and living arrangements (living alone and living with others) as covariates. More information about linear regression models is shown in Annex 2 in Multimedia Appendix 1.

[b]Not applicable.

[c]MMSE: Mini-Mental State Exam.

[d]Robust linear regression.

[e]QoL-AD: Quality of Life-Alzheimer's Disease Scale.

[f]EuroQoL-5D-3L: European Quality of Life 5 Dimensions 3 Levels.

[g]GDS: Geriatric Depression Scale.

[h]PSS: Perceived Stress Scale.

[i]As the residuals of the model were not normal, we transformed the dependent variable in its logarithmic form.

Health Care and Social Support Services Access and Utilization and Informative-Related Uses of ICTs During the COVID-19 Outbreak

Of 148 participants, 39 (26.4%) reported accessing extreme and 32 (21.6%) reported accessing too much COVID-19 information. The most frequent ICT used to access COVID-19 information was mainly the television (134/147, 91.2%), and most participants were also informed through family and friends (120/148, 81.1%). Furthermore, only 46 (30.7%) of 150 participants did not contact health or social services (Table 5; note: significant results with $P<.001$).

Table 5. Health care and social support service access and utilization and informative-related uses of ICTs[a] during the COVID-19 outbreak and differences between living alone and with others, and high and low technophilia.

Characteristic	Total participants (N=151)	Living alone (n=36)	Living with others (n=115)	Chi-square (*df*)	*P* value	High technophilia (n=80)	Low technophilia (n=65)	Chi-square (*df*)	*P* value
COVID-19 information access, n (%)									
None	3 (2.0)	1 (2.9)	2 (1.8)	2.42 (4)	.66	0	1 (1.6)	4.21 (4)	.38
Too little	33 (22.3)	8 (23.5)	25 (21.9)	2.42 (4)	.66	22 (27.5)	10 (15.9)	4.21 (4)	.38
Moderate	41 (27.7)	12 (35.3)	29 (25.4)	2.42 (4)	.66	20 (25.0)	19 (30.2)	4.21 (4)	.38
Too much	32 (21.6)	7 (20.6)	25 (21.9)	2.42 (4)	.66	16 (20.0)	16 (25.4)	4.21 (4)	.38
Extreme	39 (26.4)	6 (17.6)	33 (28.9)	2.42 (4)	.66	22 (27.5)	17 (27.0)	4.21 (4)	.38
COVID-19 information source, n (%)									
Family and friends	120 (81.1)	29 (85.3)	91 (79.8)	0.51 (1)	.48	64 (80.0)	55 (87.3)	1.35 (1)	.25
Television	134 (91.2)	32 (94.1)	102 (90.3)	0.48 (1)	.49	70 (88.6)	61 (96.8)	3.31 (1)	.07
Smartphone	56 (38.1)	13 (37.1)	43 (38.4)	0.02 (1)	.89	34 (42.5)	22 (35.5)	0.72 (1)	.40
Tablet	12 (8.2)	1 (2.9)	11 (9.8)	1.73 (1)	.19	10 (12.7)	2 (3.2)	4.08 (1)	.04
Laptop	10 (6.8)	1 (2.9)	9 (8.0)	1.06 (1)	.30	6 (7.7)	4 (6.3)	0.10 (1)	.76
Newspaper	12 (8.2)	0	12 (10.6)	3.93 (1)	.05	7 (8.9)	5 (7.9)	0.04 (1)	.84
Digital media	71 (49.0)	16 (47.1)	55 (49.5)	0.07 (1)	.80	41 (52.6)	30 (48.4)	0.24 (1)	.62
Radio	37 (24.5)	10 (29.4)	27 (24.1)	0.39 (1)	.53	18 (22.8)	19 (30.6)	1.11 (1)	.29
Resources contacted, n (%)									
None	46 (30.7)	10 (27.8)	36 (31.6)	0.19(1)	.67	21 (26.3)	21 (32.8)	0.74 (1)	.39
Health services	88 (58.3)	17 (47.2)	71 (61.7)	2.38 (1)	.12	46 (57.5)	38 (58.5)	0.01 (1)	.91
COVID-19 services	5 (3.3)	1 (2.8)	4 (3.5)	0.42 (1)	.84	4 (5.0)	1 (1.5)	1.29 (1)	.26
Emergency services	10 (6.6)	3 (8.3)	7 (6.1)	0.22 (1)	.64	5 (6.3)	5 (7.7)	0.12 (1)	.73
Social service non-government organization (NGO)	5 (3.3)	3 (8.3)	2 (1.8)	3.68 (1)	.06	4 (5.0)	1 (1.6)	1.25 (1)	.26

[a]ICT: information and communications technology.

Cognitive-, Entertainment-, and Socialization-related Uses of ICTs During the COVID-19 Outbreak

Although most of the participants (46/151, 30.7%) preferred paper-based memory exercises, the most frequent ICT used for cognition was the television (16/151, 10.7%). The most used ICTs for entertainment were the television (138/151, 92%), followed by the smartphone (60/151, 40%), and for socialization, the smartphone (75/151, 50.3%). Detailed information is given in Table 6 (note: significant results with $P<.001$).

Table 6. Cognitive-, entertainment-, and socialization-related uses of ICTs[a] during the COVID-19 outbreak and differences between living alone and living with others, and high and low technophilia.

Activity category	Total participants (N=151)	Living alone (n=36)	Living with others (n=115)	Chi-square (*df*)	*P* value	High technophilia (n=80)	Low technophilia (n=65)	Chi-square (*df*)	*P* value
Cognitive, n (%)									
Paper	46 (30.7)	14 (40.0)	32 (27.8)	7.52 (5)	.19	23 (28.7)	21 (32.8)	5.44 (5)	.36
Smartphone	3 (2.0)	0	3 (2.6)	7.52 (5)	.19	3 (3.8)	0	5.44 (5)	.36
Tablet	7 (4.7)	1 (2.9)	6 (5.2)	7.52 (5)	.19	3 (3.8)	4 (6.3)	5.44 (5)	.36
Laptop	1 (0.7)	1 (2.9)	0	7.52 (5)	.19	0	1 (1.6)	5.44 (5)	.36
Television	16 (10.7)	5 (14.3)	11 (9.6)	7.52 (5)	.19	11 (13.8)	5 (7.8)	5.44 (5)	.36
Entertainment, n (%)									
Smartphone	60 (40.0)	13 (37.1)	47 (40.9)	0.16 (1)	.69	37 (46.3)	23 (35.9)	1.56 (1)	.21
Tablet	18 (12.0)	1 (2.9)	17 (14.8)	3.61 (1)	.07	11 (13.8)	7 (10.9)	0.26 (1)	.61
Laptop	20 (13.3)	3 (8.6)	17 (14.8)	0.90 (1)	.41	12 (15.0)	8 (12.5)	0.19 (1)	.67
Television	138 (92.0)	31 (88.6)	107 (93.0)	0.73 (1)	.39	74 (92.5)	59 (92.2)	0.01 (1)	.94
Socialization, n (%)									
Home visits	87 (58.8)	29 (82.9)	58 (51.3)	10.97 (1)	<.001	46 (57.5)	40 (63.5)	0.53 (1)	.47
Smartphone	75 (50.3)	16 (45.7)	59 (51.8)	0.39 (1)	.53	47 (58.8)	27 (42.2)	3.90 (1)	.05
Tablet	10 (6.8)	0	10 (8.8)	3.32 (1)	.12	7 (8.9)	3 (4.7)	0.95 (1)	.33
Laptop	5 (3.4)	0	5 (4.4)	1.59 (1)	.59	3 (3.8)	2 (3.1)	0.04 (1)	.84
Television	6 (4.0)	0	6 (5.3)	1.91 (1)	.34	5 (6.3)	1 (1.6)	1.96 (1)	.23

[a]ICT: information and communications technology.

Discussion

Principal Findings

This cohort study was conducted to understand the impact of restrictive measures in community-dwelling older adults with MCI and MD during the first COVID-19 outbreak.

Our findings show that the first months of the outbreak did not significantly impact the cognition, quality of life, perceived health status, and depression of our study population when making comparisons with baseline assessments prior to the outbreak. Change in living arrangements had no influence on these variables either. Living alone and technophilia were not associated with mental health–related variables after correction for multiple comparisons. However, being alone was nominally associated with self-perceived fear and depression, and higher technophilia with better quality of life, less boredom, perceived stress, and depression but also less calmness. Overall, health care and social support service access and utilization were high. The most used ICTs during the COVID-19 outbreak were the television for informative-, cognitive-, and entertainment-related uses and the smartphone for socialization.

Comparison With Prior Work

To the best of our knowledge, few studies have addressed the consequences of the COVID-19 outbreak on the cognition of the elderly, and the use of technologies during this ongoing societal change.

Several studies have shown that quarantine measures have changed the behaviors and lifestyle of older people with cognitive decline [42], although in some cases, these changes have been less important than expected [43]. Lifestyle changes can increase the risk of dementia and cause cognitive impairment. However, our study showed that in our sample, the first stages of the COVID-19 outbreak did not cause significant cognitive decline in comparison with a previous assessment. This result could be explained by 2 reasons. On the one hand, the evaluation was carried out on an average of 70 days after the start of the home confinement restrictions and, likely, this time was not enough to influence cognitive decline. On the other hand, the data of our study show that most of the participants maintained an active lifestyle and used new technologies for cognitive stimulation, information access, leisure, and social connectedness. This combination of healthy lifestyle factors and opportunities for cognitive and social stimulation has proved to be important in reducing the risk of cognitive decline [44].

Regarding the lack of differences in the quality of life and perceived health status before and during the outbreak in people with MCI/MD, a similar conclusion was reached by another cohort study in a similar population in Spain using the EuroQoL-5D-3L [45]. Other studies on quality of life in different population groups during the pandemic have also not found a perception of poor quality of life in the elderly [46]. Our results could be explained because our sample has continued to carry out cognitive, entertainment, and social activities, which are associated with the definition of a good quality of life by

the elderly [47]. Another factor that can lead to a good quality of life is maintaining well-being and good sleep quality. Poor sleep quality is known to have a significant impact on lower levels of life satisfaction and mood [48].

Regarding mental health, longitudinal studies have established that it is has been affected by the pandemic [49-51], but frequently, the impact has been higher in the younger population and those more economically vulnerable [51,52]. Likewise, another longitudinal study in Spain with a sample of older adults with dementia or cognitive impairment found an increase in depressive and anxious symptoms after the confinement [53]. Furthermore, attention has been drawn to the possible harmful effects of the excessively dramatic presentation of the consequences of the restrictions due to the pandemic, which are based mainly on survey studies, in many cases carried out without the required rigor [54]. However, the results are mixed, and, for example, some longitudinal studies in Spain [55] and Greece [56] did not find differences regarding depression when comparing the period before and after confinement, and another Dutch community study did not find an increase in depressive symptoms in the general population [57]. In a cross-sectional study comparing populations over and under the age of 60 years during the peak of the COVID-19 pandemic in Spain, the elderly did not demonstrate special vulnerability to acute stress and no sex differences were found. This study hypothesized a greater resilience in the elderly due to the economic and social difficulties experienced throughout their lives during the Spanish post–Civil War period (1939-1960), increasing their ability to cope with stress and face the pandemic resiliently [52]. The results of this study in which no differences were found in depressive symptoms could also be explained by the specific characteristics of the study sample, which present MCI and mainly maintain autonomy to live independently. In addition, the evaluations were conducted when the stricter measures of the first lockdown in Spain were being brought down. Moreover, the small incidence of the COVID-19 virus at the time of assessment could have influenced the results (infections less than 1%). These results show the complexity of the effects of the pandemic, highlighting the need for more longitudinal studies in different populations to evaluate the effects of the social restrictions and the pandemic.

Another factor to consider was whether living alone during COVID-19 confinement was associated with a higher prevalence of depression. Although several studies have found a significant association between depression and living alone during the pandemic [42,58], others have shown otherwise [59,60]. Our results are in line with the second ones, showing no significant differences in GDS scores. However, more studies are necessary to determine whether living alone is a risk factor for depression during the pandemic. The results are in line with our previous findings [61] that associated living alone with worse mental health at the beginning of the outbreak (without correcting for multiple comparisons). The way our sample used the ICTs, through online communication, remote social interactions, or video calls, could have been useful to address social isolation during the pandemic [62].

Regarding technophilia, our study did not find an association between a better attitude toward technology and better mental health. In line with these results, a multicenter study conducted in Norway, the United Kingdom, the United States, and Australia also found no change in loneliness and the quality of life in adults over 70 years who used ICTs to maintain social contact during the COVID outbreak [63]. However, we showed in a previous study based on the TV-AssistDem RCT how technology could be useful to maintain cognitive activities [61], and more studies need to clarify whether the evidence supports the recommendations on interventions that may improve the knowledge of ICTs and are related with the use of technology to maintain social connections and cognitive activities [64].

Limitations

A main limitation of this study was changing the interview administration from face-to-face before the outbreak to telephonic during the outbreak. Interviews were performed by the same professionals in both cases to reduce this possible bias, the measures were rescaled accordingly, and the validated phone versions of the tests were used. In addition, the interview had a mean duration of 50.14 minutes, which could cause fatigue in this population and alter their performance.

Another limitation is that the sample came from 2 RCTs and the participants who agreed to participate in the RCTs may have special characteristics that make them not representative of the general population. Furthermore, the sample was from only 1 center in Andalusia. However, it is a larger sample than in other studies carried out in Spain to date with this type of population.

Moreover, 15 caregivers answered on behalf of patients whose ability to answer for themselves was compromised. Although they did not respond to the questionnaires that evaluated the main variables, their answers may have interfered with the results.

The Bonferroni correction for multiple comparisons is conservative and could have increased type II errors.

Some studies have pointed out that the effects of changes during the lockdown may be temporary compared to long-lasting ones. Therefore, future effects will need to be explored as it is possible that once the lockdown is over, many people may not return to their “normal routine” as before the pandemic and will continue to avoid face-to-face activities, especially those regarding social and physical activity due to the fear of the contagion [42].

Conclusions

During the COVID-19 outbreak, governments’ restrictive measures demonstrated being effective in viral spread prevention. Although these restrictions have had negative effects on health and well-being and have changed lifestyles worldwide, our study showed how a presumably vulnerable population has shown more resilience to restrictive measures than expected. The people with MCI/MD did not experience a significant decline in cognition, quality of life, perceived health status, or depression during the period of the COVID-19 outbreak. The study also showed that being alone and a negative attitude toward technology are not associated with worse mental health after correcting for multiple comparisons. In addition, the data were collected over a short period, and further research is needed to explore whether maintaining restrictive measures for longer

influences a worsening of cognitive abilities, quality of life, perceived health status, or depression and which factors increase the risk of poor mental health in this population. They reported overall well-being, maintained sleep quality, and presented moderate perceived stress. This could be related to the fact that our sample continued participating in daily activities, which plays a crucial role in enhancing and maintaining cognition [65,66], just like keeping social interaction. The use of ICTs was essential to carry out these activities during the restrictions. Television was the most widely used ICT for informational-, cognitive-, and entertainment-related uses, and the smartphone for socialization.

Authors' Contributions

The authorship of this manuscript follows the International Committee of Medical Journal Editors standards. EDP, JMGC, AVN, CGSL, FMC, and JGP made substantial contributions to the conception and design of the work; EDP, JMGC, GGP, AVN, EVM, MG, MQ, and PBF acquired the data; JGP analyzed and interpreted the data; and EDP, JMGC, and JGP drafted the work. All authors revised it critically for important intellectual content and gave final approval of the version to be published.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Model of the questionnaire used.
[DOCX File , 60 KB - jmir_v24i2e30598_app1.docx]

References

1. WHO Director-General's Opening Remarks at the Media Briefing on COVID-19 - 11 March 2020. URL: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020 [accessed 2022-02-16]
2. BOE. Real Decreto 463/2020, de 14 de marzo, por el que se declara el estado de alarma para la gestión de la situación de crisis sanitaria ocasionada por el COVID-19. 2021. URL: https://www.boe.es/buscar/act.php?id=BOE-A-2020-3692 [accessed 2022-02-16]
3. Santesmasses D, Castro JP, Zenin AA, Shindyapina AV, Gerashchenko MV, Zhang B, et al. COVID-19 is an emergent disease of aging. Aging Cell 2020 Oct;19(10):e13230 [FREE Full text] [doi: 10.1111/acel.13230] [Medline: 33006233]
4. Centro de Coordinación de Alertas y Emergencias Sanitarias. Actualización nº 325. Enfermedad por el coronavirus (COVID-19). 04.03.2021. URL: https://www.mscbs.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov/documentos/Actualizacion_325_COVID-19.pdf [accessed 2022-02-16]
5. Hwang J, Kim J, Park J, Chang MC, Park D. Neurological diseases as mortality predictive factors for patients with COVID-19: a retrospective cohort study. Neurol Sci 2020 Sep 08;41(9):2317-2324 [FREE Full text] [doi: 10.1007/s10072-020-04541-z] [Medline: 32643133]
6. Bianchetti A, Rozzini R, Guerini F, Boffelli S, Ranieri P, Minelli G, et al. Clinical presentation of COVID19 in dementia patients. J Nutr Health Aging 2020 May 15;24(6):560-562 [FREE Full text] [doi: 10.1007/s12603-020-1389-1] [Medline: 32510106]
7. Aggarwal G, Lippi G, Michael Henry B. Cerebrovascular disease is associated with an increased disease severity in patients with coronavirus disease 2019 (COVID-19): a pooled analysis of published literature. Int J Stroke 2020 Jun;15(4):385-389. [doi: 10.1177/1747493020921664] [Medline: 32310015]
8. Atkins JL, Masoli JAH, Delgado J, Pilling LC, Kuo CL, Kuchel GA, et al. Preexisting comorbidities predicting COVID-19 and mortality in the UK Biobank community cohort. J Gerontol A Biol Sci Med Sci 2020 Oct 15;75(11):2224-2230 [FREE Full text] [doi: 10.1093/gerona/glaa183] [Medline: 32687551]
9. Sofi F, Valecchi D, Bacci D, Abbate R, Gensini G, Casini A, et al. Physical activity and risk of cognitive decline: a meta-analysis of prospective studies. J Intern Med 2011 Jan;269(1):107-117 [FREE Full text] [doi: 10.1111/j.1365-2796.2010.02281.x] [Medline: 20831630]
10. Bieber A, Nguyen N, Meyer G, Stephan A. Influences on the access to and use of formal community care by people with dementia and their informal caregivers: a scoping review. BMC Health Serv Res 2019 Feb 01;19(1):88 [FREE Full text] [doi: 10.1186/s12913-018-3825-z] [Medline: 30709345]
11. Kuiper JS, Zuidersma M, Oude Voshaar RC, Zuidema SU, van den Heuvel ER, Stolk RP, et al. Social relationships and risk of dementia: a systematic review and meta-analysis of longitudinal cohort studies. Ageing Res Rev 2015 Jul;22:39-57. [doi: 10.1016/j.arr.2015.04.006] [Medline: 25956016]
12. Sommerlad A, Ruegger J, Singh-Manoux A, Lewis G, Livingston G. Marriage and risk of dementia: systematic review and meta-analysis of observational studies. J Neurol Neurosurg Psychiatry 2018 Mar 28;89(3):231-238 [FREE Full text] [doi: 10.1136/jnnp-2017-316274] [Medline: 29183957]

13. Livingston G, Sommerlad A, Orgeta V, Costafreda SG, Huntley J, Ames D, et al. Dementia prevention, intervention, and care. Lancet 2017 Dec;390(10113):2673-2734. [doi: 10.1016/s0140-6736(17)31363-6]
14. Tomaka J, Thompson S, Palacios R. The relation of social isolation, loneliness, and social support to disease outcomes among the elderly. J Aging Health 2006 Jun 30;18(3):359-384. [doi: 10.1177/0898264305280993] [Medline: 16648391]
15. Domènech-Abella J, Lara E, Rubio-Valera M, Olaya B, Moneta MV, Rico-Uribe LA, et al. Loneliness and depression in the elderly: the role of social network. Soc Psychiatry Psychiatr Epidemiol 2017 Apr 2;52(4):381-390. [doi: 10.1007/s00127-017-1339-3] [Medline: 28154893]
16. Wilson RS, Krueger KR, Arnold SE, Schneider JA, Kelly JF, Barnes LL, et al. Loneliness and risk of Alzheimer disease. Arch Gen Psychiatry 2007 Feb 01;64(2):234-240. [doi: 10.1001/archpsyc.64.2.234] [Medline: 17283291]
17. Courtin E, Knapp M. Social isolation, loneliness and health in old age: a scoping review. Health Soc Care Community 2017 May 28;25(3):799-812. [doi: 10.1111/hsc.12311] [Medline: 26712585]
18. Hwang T, Rabheru K, Peisah C, Reichman W, Ikeda M. Loneliness and social isolation during the COVID-19 pandemic. Int Psychogeriatr 2020 May 26;32(10):1217-1220. [doi: 10.1017/s1041610220000988]
19. Savva GM, Zaccai J, Matthews FE, Davidson JE, McKeith I, Brayne C, Medical Research Council Cognitive Function Ageing Study. Prevalence, correlates and course of behavioural and psychological symptoms of dementia in the population. Br J Psychiatry 2009 Mar 02;194(3):212-219. [doi: 10.1192/bjp.bp.108.049619] [Medline: 19252147]
20. Valenzuela MJ, Sachdev P. Brain reserve and dementia: a systematic review. Psychol Med 2005 Oct 06;36(4):441-454. [doi: 10.1017/s0033291705006264]
21. Cotterell N, Buffel T, Phillipson C. Preventing social isolation in older people. Maturitas 2018 Jul;113:80-84. [doi: 10.1016/j.maturitas.2018.04.014] [Medline: 29903652]
22. Smith AC, Thomas E, Snoswell CL, Haydon H, Mehrotra A, Clemensen J, et al. Telehealth for global emergencies: implications for coronavirus disease 2019 (COVID-19). J Telemed Telecare 2020 Mar 20;26(5):309-313. [doi: 10.1177/1357633x20916567]
23. Goodman-Casanova J, Dura-Perez E, Guerrero-Pertiñez G, Barnestein-Fonseca P, Guzman-Parra J, Vega-Nuñez A, et al. Cognitive outcomes during COVID-19 confinement among older people and their caregivers using technologies for dementia: protocol for an observational cohort study. JMIR Res Protoc 2021 May 18;10(5):e26431. [doi: 10.2196/26431]
24. Anderberg P, Barnestein-Fonseca P, Guzman-Parra J, Garolera M, Quintana M, Mayoral-Cleries F, et al. The effects of the digital platform Support Monitoring and Reminder Technology for Mild Dementia (SMART4MD) for people with mild cognitive impairment and their informal carers: protocol for a pilot randomized controlled trial. JMIR Res Protoc 2019 Jun 21;8(6):e13711 [FREE Full text] [doi: 10.2196/13711] [Medline: 31228177]
25. Goodman-Casanova JM, Guzmán-Parra J, Guerrero G, Vera E, Barnestein-Fonseca P, Cortellessa G, et al. TV-based assistive integrated service to support European adults living with mild dementia or mild cognitive impairment (TV-AssistDem): study protocol for a multicentre randomized controlled trial. BMC Geriatr 2019 Sep 06;19(1):247 [FREE Full text] [doi: 10.1186/s12877-019-1267-z] [Medline: 31492113]
26. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6]
27. O'Bryant SE, Humphreys JD, Smith GE, Ivnik RJ, Graff-Radford NR, Petersen RC, et al. Detecting dementia with the mini-mental state examination in highly educated individuals. Arch Neurol 2008 Jul 14;65(7):963-967 [FREE Full text] [doi: 10.1001/archneur.65.7.963] [Medline: 18625866]
28. Roccaforte WH, Burke WJ, Bayer BL, Wengel SP. Validation of a telephone version of the mini-mental state examination. J Am Geriatr Soc 1992 Jul 27;40(7):697-702. [doi: 10.1111/j.1532-5415.1992.tb01962.x] [Medline: 1607586]
29. Gibbons L, Mccurry S, Teri L. Quality of life in Alzheimer's disease: patient and caregiver reports. J Mental Health Aging 1999;5:32.
30. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec 9;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
31. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med 2001 Jul 08;33(5):337-343. [doi: 10.3109/07853890109002087] [Medline: 11491192]
32. Brazier J, Connell J, Papaioannou D, Mukuria C, Mulhern B, Peasgood T, et al. A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures. Health Technol Assess 2014 May;18(34):vii-viii, xiii [FREE Full text] [doi: 10.3310/hta18340] [Medline: 24857402]
33. Kunz S. Psychometric properties of the EQ-5D in a study of people with mild to moderate dementia. Qual Life Res 2010 Apr 10;19(3):425-434. [doi: 10.1007/s11136-010-9600-1] [Medline: 20146008]
34. Yesavage JA, Brink T, Rose TL, Lum O, Huang V, Adey M, et al. Development and validation of a geriatric depression screening scale: a preliminary report. J Psychiatr Res 1982 Jan;17(1):37-49. [doi: 10.1016/0022-3956(82)90033-4]
35. Carrete P, Augustovski F, Gimpel N, Fernandez S, Di Paolo R, Schaffer I, et al. Validation of a telephone-administered geriatric depression scale in a Hispanic elderly population. J Gen Intern Med 2001 Jul;16(7):446-450 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016007446.x] [Medline: 11520381]

36. Anderberg P, Eivazzadeh S, Berglund JS. A novel instrument for measuring older people's attitudes toward technology (TechPH): development and validation. J Med Internet Res 2019 May 23;21(5):e13951 [FREE Full text] [doi: 10.2196/13951] [Medline: 31124467]
37. Remor E. Psychometric properties of a European Spanish version of the Perceived Stress Scale (PSS). Span J Psychol 2006 May 10;9(1):86-93. [doi: 10.1017/s1138741600006004] [Medline: 16673626]
38. Campo-Arias A, Pedrozo-Cortés MJ, Pedrozo-Pupo JC. Pandemic-related Perceived Stress Scale of COVID-19: an exploration of online psychometric performance. Revista Colombiana de Psiquiatría (English ed) 2020 Oct;49(4):229-230. [doi: 10.1016/j.rcpeng.2020.05.001]
39. Pedrozo-Pupo JC, Pedrozo-Cortés MJ, Campo-Arias A. Perceived stress associated with COVID-19 epidemic in Colombia: an online survey. Cad Saude Publica 2020;36(5):e00090520 [FREE Full text] [doi: 10.1590/0102-311x00090520] [Medline: 32490918]
40. Gordon M. Assess Notes: Nursing Assessment & Diagnostic Reasoning. Philadelphia, PA: F.A. Davis; 2008.
41. R Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2014.
42. Di Santo SG, Franchini F, Filiputti B, Martone A, Sannino S. The effects of COVID-19 and quarantine measures on the lifestyles and mental health of people over 60 at increased risk of dementia. Front Psychiatry 2020 Oct 14;11:578628 [FREE Full text] [doi: 10.3389/fpsyt.2020.578628] [Medline: 33173523]
43. Lehtisalo J, Palmer K, Mangialasche F, Solomon A, Kivipelto M, Ngandu T. Changes in lifestyle, behaviors, and risk factors for cognitive impairment in older persons during the first wave of the coronavirus disease 2019 pandemic in Finland: results from the FINGER Study. Front Psychiatry 2021;12:624125 [FREE Full text] [doi: 10.3389/fpsyt.2021.624125] [Medline: 33643095]
44. Kivipelto M, Mangialasche F, Ngandu T. Lifestyle interventions to prevent cognitive impairment, dementia and Alzheimer disease. Nat Rev Neurol 2018 Nov 5;14(11):653-666. [doi: 10.1038/s41582-018-0070-3] [Medline: 30291317]
45. Lara B, Carnes A, Dakterzada F, Benitez I, Piñol-Ripoll G. Neuropsychiatric symptoms and quality of life in Spanish patients with Alzheimer's disease during the COVID-19 lockdown. Eur J Neurol 2020 Sep 24;27(9):1744-1747 [FREE Full text] [doi: 10.1111/ene.14339] [Medline: 32449791]
46. Bidzan-Bluma I, Bidzan M, Jurek P, Bidzan L, Knietzsch J, Stueck M, et al. A Polish and German population study of quality of life, well-being, and life satisfaction in older adults during the COVID-19 pandemic. Front Psychiatry 2020 Nov 17;11:585813 [FREE Full text] [doi: 10.3389/fpsyt.2020.585813] [Medline: 33281646]
47. Xavier FMF, Ferraz MPT, Marc N, Escosteguy NU, Moriguchi EH. Elderly people's definition of quality of life. Braz J Psychiatry 2003 Mar;25(1):31-39 [FREE Full text] [doi: 10.1590/s1516-44462003000100007] [Medline: 12975677]
48. Zhi T, Sun X, Li S, Wang Q, Cai J, Li L, et al. Associations of sleep duration and sleep quality with life satisfaction in elderly Chinese: the mediating role of depression. Arch Gerontol Geriatr 2016 Jul;65:211-217. [doi: 10.1016/j.archger.2016.03.023] [Medline: 27100684]
49. Pierce M, Hope H, Ford T, Hatch S, Hotopf M, John A, et al. Mental health before and during the COVID-19 pandemic: a longitudinal probability sample survey of the UK population. Lancet Psychiatry 2020 Oct;7(10):883-892. [doi: 10.1016/s2215-0366(20)30308-4]
50. Wong SYS, Zhang D, Sit RWS, Yip BHK, Chung RY, Wong CKM, et al. Impact of COVID-19 on loneliness, mental health, and health service utilisation: a prospective cohort study of older adults with multimorbidity in primary care. Br J Gen Pract 2020 Sep 28;70(700):e817-e824. [doi: 10.3399/bjgp20x713021]
51. Fancourt D, Steptoe A, Bu F. Trajectories of anxiety and depressive symptoms during enforced isolation due to COVID-19 in England: a longitudinal observational study. Lancet Psychiatry 2021 Feb;8(2):141-149. [doi: 10.1016/s2215-0366(20)30482-x]
52. García-Fernández L, Romero-Ferreiro V, López-Roldán PD, Padilla S, Rodriguez-Jimenez R. Mental health in elderly Spanish people in times of COVID-19 outbreak. Am J Geriatr Psychiatry 2020 Oct;28(10):1040-1045 [FREE Full text] [doi: 10.1016/j.jagp.2020.06.027] [Medline: 32718855]
53. Barguilla A, Fernández-Lebrero A, Estragués-Gázquez I, García-Escobar G, Navalpotro-Gómez I, Manero RM, et al. Effects of COVID-19 pandemic confinement in patients with cognitive impairment. Front Neurol 2020 Nov 24;11:589901 [FREE Full text] [doi: 10.3389/fneur.2020.589901] [Medline: 33329337]
54. Hawton K, Marzano L, Fraser L, Hawley M, Harris-Skillman E, Lainez YX. Reporting on suicidal behaviour and COVID-19: need for caution. Lancet Psychiatry 2021 Jan;8(1):15-17. [doi: 10.1016/s2215-0366(20)30484-3]
55. Fernández-Abascal EG, Martín-Díaz MD. Longitudinal study on affect, psychological well-being, depression, mental and physical health, prior to and during the COVID-19 pandemic in Spain. Pers Individ Dif 2021 Apr;172:110591 [FREE Full text] [doi: 10.1016/j.paid.2020.110591] [Medline: 33518870]
56. Tsatali M, Moraitou D, Poptsi E, Sia E, Agogiatou C, Gialaouzidis M, et al. Are there any cognitive and behavioral changes potentially related to quarantine due to the COVID-19 pandemic in people with mild cognitive impairment and AD dementia? A longitudinal study. Brain Sci 2021 Aug 31;11(9):1165 [FREE Full text] [doi: 10.3390/brainsci11091165] [Medline: 34573186]

57. van der Velden PG, Hyland P, Contino C, von Gaudecker H, Muffels R, Das M. Anxiety and depression symptoms, the recovery from symptoms, and loneliness before and after the COVID-19 outbreak among the general population: findings from a Dutch population-based longitudinal study. PLoS One 2021 Jan 7;16(1):e0245057 [FREE Full text] [doi: 10.1371/journal.pone.0245057] [Medline: 33411843]
58. Aragón I, Flores I, Dorman G, Rojas G, Sierra Sanjurjo N, O'Neill S. Quality of life, mood, and cognitive performance in older adults with cognitive impairment during the first wave of COVID 19 in Argentina. Int J Geriatr Psychiatry 2022 Jan 16;37(1):Epub [FREE Full text] [doi: 10.1002/gps.5650] [Medline: 34755917]
59. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. Lancet 2020 Mar;395(10227):912-920. [doi: 10.1016/s0140-6736(20)30460-8]
60. van den Besselaar JH, MacNeil Vroomen JL, Buurman BM, Hertogh CM, Huisman M, Kok AA, et al. Symptoms of depression, anxiety, and perceived mastery in older adults before and during the COVID-19 pandemic: results from the Longitudinal Aging Study Amsterdam. J Psychosom Res 2021 Dec;151:110656 [FREE Full text] [doi: 10.1016/j.jpsychores.2021.110656] [Medline: 34741872]
61. Goodman-Casanova J, Dura-Perez E, Guzman-Parra J, Cuesta-Vargas A, Mayoral-Cleries F. Telehealth home support during COVID-19 confinement for community-dwelling older adults with mild cognitive impairment or mild dementia: survey study. In: J Med Internet Res. Telehealth home support during COVID-19 confinement for community-dwelling older adults with mild cognitive impairment or mild dementia: Survey study. Vol. 22, Journal of Medical Internet Research. JMIR Publications; May 22, 2020:e19434.
62. Arpino B, Pasqualini M, Bordone V, Solé-Auró A. Older people's nonphysical contacts and depression during the COVID-19 lockdown. Gerontologist 2021 Feb 23;61(2):176-186 [FREE Full text] [doi: 10.1093/geront/gnaa144] [Medline: 32977334]
63. Bonsaksen T, Thygesen H, Leung J, Ruffolo M, Schoultz M, Price D, et al. Video-based communication and its association with loneliness, mental health and quality of life among older people during the COVID-19 outbreak. Int J Environ Res Public Health 2021 Jun 10;18(12):6284 [FREE Full text] [doi: 10.3390/ijerph18126284] [Medline: 34200670]
64. Sepúlveda-Loyola W, Rodríguez-Sánchez I, Pérez-Rodríguez P, Ganz F, Torralba R, Oliveira DV, et al. Impact of social isolation due to COVID-19 on health in older people: mental and physical effects and recommendations. J Nutr Health Aging 2020 Sep 25;24(9):938-947. [doi: 10.1007/s12603-020-1500-7]
65. Etnier JL, Drollette ES, Slutsky AB. Physical activity and cognition: a narrative review of the evidence for older adults. Psychol Sport Exerc 2019 May;42:156-166. [doi: 10.1016/j.psychsport.2018.12.006]
66. Singh-Manoux A, Richards M, Marmot M. Leisure activities and cognitive function in middle age: evidence from the Whitehall II study. J Epidemiol Community Health 2003 Nov 01;57(11):907-913 [FREE Full text] [doi: 10.1136/jech.57.11.907] [Medline: 14600119]

Abbreviations

EuroQoL-5D-3L: European Quality of Life 5 Dimensions 3 Levels
GDS: Geriatric Depression Scale
HRQoL: health-related quality of life
ICT: information and communications technology
MCI: mild cognitive impairment
MD: mild dementia
MMSE: Mini-Mental State Examination
PSS: Perceived Stress Scale
QoL-AD: Quality of Life-Alzheimer’s Disease Scale
RCT: randomized clinical trial
SMART4MD: Support, Monitoring and Reminder Technology for Mild Dementia
TechPH: Instrument for Measuring Older People’s Attitudes Toward Technology
TV-AssistDem: TV-Based Assistive Integrated Service to Support European Adults Living with Dementia
VAS: visual analog scale

Edited by C Basch; submitted 21.05.21; peer-reviewed by C Sitges, MS Hosseini, A Garcia; comments to author 10.07.21; revised version received 13.10.21; accepted 23.12.21; published 22.02.22.

Please cite as:
Dura-Perez E, Goodman-Casanova JM, Vega-Nuñez A, Guerrero-Pertiñez G, Varela-Moreno E, Garolera M, Quintana M, Cuesta-Vargas AI, Barnestein-Fonseca P, Gómez Sánchez-Lafuente C, Mayoral-Cleries F, Guzman-Parra J
The Impact of COVID-19 Confinement on Cognition and Mental Health and Technology Use Among Socially Vulnerable Older People: Retrospective Cohort Study
J Med Internet Res 2022;24(2):e30598
URL: https://www.jmir.org/2022/2/e30598
doi:10.2196/30598
PMID:35049505

Original Paper

Telehealth Before and During the COVID-19 Pandemic: Analysis of Health Care Workers' Opinions

Pascal Nitiema[1], MD, MPH, MSc

Division of Management Information Systems, Price College of Business, University of Oklahoma, Norman, OK, United States

Corresponding Author:
Pascal Nitiema, MD, MPH, MSc
Division of Management Information Systems
Price College of Business
University of Oklahoma
307 W Brooks Suite 307
Norman, OK, 73019
United States
Phone: 1 4053255721
Email: pascal.nitiema-1@ou.edu

Abstract

Background: The COVID-19 pandemic and the lockdowns for controlling the spread of infection have led to a surge in telehealth adoption by many health care organizations. It is unclear how this pandemic has impacted health professionals' view about telehealth. The analysis of textual data, such as comments posted on a discussion forum, can uncover information that may not be captured by a structured survey.

Objective: This study aims to examine the opinions of health care workers about telehealth services during the time frame of March 2013-December 2020.

Methods: Comments about telehealth posted by health care workers from at least 46 countries were collected from an online discussion forum dedicated to health professionals. The analysis included the computation of sentiment scores from the textual data and the use of structural topic modeling to identify the topics of discussions as well as the factors that may be associated with the prevalence of these topics.

Results: The analysis of the comments revealed positive opinions about the perceived benefits of telehealth services before and during the pandemic, especially the ability to reach patients who cannot come to the health facility for diverse reasons. However, opinions about these benefits were less positive during the pandemic compared to the prepandemic period. Specific issues raised during the pandemic included technical difficulties encountered during telehealth sessions and the inability to perform certain care routines through telehealth platforms. Although comments on the quality of care provided through telehealth were associated with a negative sentiment score overall, the average score was less negative during the pandemic compared to the prepandemic period, signaling a shift in opinion about the quality of telehealth services. In addition, the analysis uncovered obstacles to the adoption of telehealth, including the absence of adequate legal dispositions for telehealth services and issues regarding the payment of these services by health insurance organizations.

Conclusions: Enhancing the adoption of telehealth services beyond the pandemic requires addressing issues related to the quality of care, payment of services, and legal dispositions for delivering these services.

(J Med Internet Res 2022;24(2):e29519) doi:10.2196/29519

KEYWORDS

telehealth; telemedicine; COVID-19; pandemic; physical examination; sentiment score; structural topic modeling; opinion; health care worker; social media; discussion

Introduction

Telehealth services have been described by the World Health Organization as having a "great potential to address some of the challenges faced by both developed and developing countries in providing accessible, cost-effective, high-quality health care services" [1]. Multiple efforts have been undertaken across countries to increase the scope and reach of these services. Although the rate of adoption of telehealth by health care

organizations in the United States has been steadily increasing, the percentage of hospitals using telehealth services is still relatively low [2,3]. Only 15% of physicians in the United States worked in health practices that used telehealth services in 2016 [3].

The COVID-19 pandemic and the lockdowns mandated by public authorities to reduce the spread of infection have led to a renewed interest from health care organizations in adopting telehealth platforms for some of their routine patient care operations. Event system theory posits that organizations are dynamic entities that can respond to events that are novel, critical (ie, the extent to which the event is deemed significant for the organization), and disruptive by altering their routine practices so that they can adapt to these events [4]. These unexpected events, because of their novel nature, usually find the organizations ill-prepared, causing the interruption of day-to-day operations and requiring changes to existing practices or the adoption of new ones [5]. Hence, it is not surprising that health care organizations have turned to telehealth services to accommodate the disruptions in patient care operations that were caused by the pandemic. Telehealth services allow long-distance patient care through telecommunication channels and thus do not require that patients and health care workers involved in the care process be physically in the same room. This feature made telehealth useful for adapting to the mandated lockdowns and the safe physical distancing measures advocated during the pandemic.

However, the successful implementation of changes in processes requires the adherence of organizational members. An unexpected event disrupts not only the routines of organizations but also the automatic cognitive processing employees use to perform their tasks [6]. Automatic cognitive processing, also called automatic information processing, is built through "repetitive training on the same task" and allows performing job-related tasks relatively swiftly and efficiently with reduced cognitive effort [7]. For instance, clinicians who have conducted physical examinations of hundreds or thousands of patients have developed heuristics and routinized practices for arriving at a diagnosis after detecting specific lung or heart sounds during auscultation. The disruption in automatic cognitive processing and the need to acquire new skills proper for the new circumstances created by a disruptive event impose undue burden on the employees and can lead to occupational stress and job dissatisfaction. However, the adoption of new practices can also be positively perceived and considered the rightful way for navigating the challenges created by the disruptive event.

The objective of this research is to examine the opinions of health care workers about telehealth services expressed before and during the COVID-19 pandemic. Comparing comments posted during the 2 time frames will illustrate the influence of the pandemic on the viewpoints of these health care workers on telehealth. The analysis of these opinions can shed light on the obstacles encountered by health care professionals using these services and help identify solutions to promote the adoption of these services.

Methods

Definition of Telehealth

The definitions of telehealth and telemedicine often vary across authors. In this manuscript, *telehealth* is defined as the umbrella term for a set of activities and services performed by health care professionals through telecommunication technologies to "support and promote long-distance clinical health care, patient and professional health-related education, public health and health administration" [8]. Hence, telehealth includes collaboration among health care workers discussing and sharing patients' information through telecommunication channels, data collection and remote monitoring of patients' health outcomes through digital wearables, and electronic transmission of prescriptions to pharmacists (e-prescribing). *Telemedicine*, a subset of telehealth, is strictly defined as the diagnosis and treatment of patients through telecommunication technologies. Nevertheless, the term *telehealth* will be used throughout the manuscript, except when analyzing or quoting health care workers' comments, to encompass all health care operations performed through telecommunication technologies.

Data Collection

The comments were collected from Medscape, an online platform that provides health care–related news to health professionals and hosts forums on health-related issues. According to the web analytics company similarweb, the online platform Medscape had more than 21 million visits per month during the summer and fall seasons of the year 2020. Most of the visits were from the United States (53%). The list of countries of the platform visitors included the United Kingdom (4%), Australia (4%), Canada (4%), India (3%), the Philippines (2%), and Brazil (2%). The keywords used to search for the comments were *telemedicine*, *telehealth*, *televisit*, *virtual visit*, *online visit*, and *video visit*. Available data on the characteristics of the commenters were collected as well, including the country of residence, occupation, and medical specialty when the commenter was a physician. Since the comments posted on the platform are publicly available and can be accessed by anyone, the institutional review board of the University of Oklahoma, USA, determined that this investigation did not require ethics committee approval.

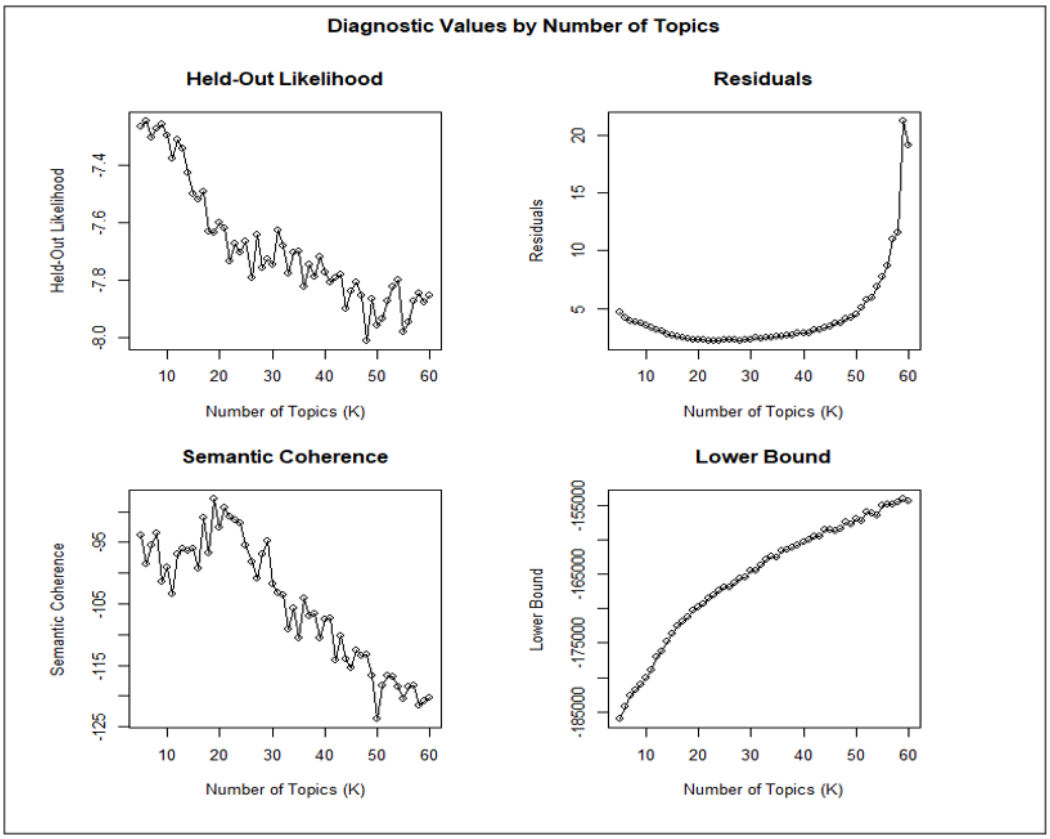
Data Analysis

The sentiment scores of the collected comments were computed using the R package *sentiment*, which considers valence shifters (eg, happy vs not happy) and (de-)amplifying words (eg, very worried vs slightly worried) [9]. A mixed-effect model with random intercepts was fitted to assess the association between sentiment scores and the following independent variables: residence of the commenter (US vs non-US), occupation (physician, nurse, health administrator, other) of the commenter, and the year in which the comment was made. The restricted maximum likelihood method was used to estimate the model parameters, and the covariance pattern of the model was set to unstructured.

Examining the prevalence of discussion topics across time frames and commenters' characteristics can shed light on the

opinions or concerns about telehealth that are specific to these time frames and these individual characteristics. Structural topic modeling (STM) was used to identify the topics discussed in the comments. STM allows the analyst to incorporate information about the characteristics (metadata) and content of the documents into the modeling process [10]. The following metadata were used as independent variables to estimate the topic prevalence in the comments: the residence (US vs non-US) and occupation (physician, nurse, health administrator, other) of the commenter, the time frame in which the comment was posted (before vs during the pandemic), and the average sentiment score of the comment. The date of March 11, 2020, the day the World Health Organization declared the COVID-19 outbreak a pandemic, was used as the cut-off date for defining the time period [11]. First, the texts of the comments were processed by removing stop words using the System for the Mechanical Analysis and Retrieval of Text (SMART) stop word list. Then, nouns in plural were converted to their singular forms. No stemming was performed. The appropriate number of topics for the corpus of comments was determined to be between 10 and 25 by examining the values for the held-out likelihood, semantic coherence, residuals, and lower bound [12]. See Figure 1. The appropriate number of topics to extract was determined by selecting the model that yielded both the highest semantic coherence and the highest exclusivity values. The number of topics found to have optimal exclusivity was 13, and thus, 13 topics were selected for modeling. Spectral initialization was used for the topic modeling, which converged after 500 iterations, and 4 single-word comments were dropped during the modeling process. The final model consisted of 13 topics, 910 documents, and a 3252-word dictionary. The correlations among the identified topics were computed with the maximum a posteriori estimates of the topic proportions. A threshold of r=0.15 for the correlation was set to determine whether 2 topics were correlated; that is, 2 topics were considered uncorrelated if their pairwise correlation was less than 0.15. STM was performed using the R package *stm* [12].

Figure 1. Diagnostic values by number of topics.



Results

Characteristics of Health Care Workers

A total of 914 comments made by 705 health care workers were collected from the forums. The dates of the comments spanned from the year 2013 through 2020, and the commenters were from at least 46 countries in 6 continents (Africa, Asia, Australia, Europe, North America, and South America). The majority of commenters were physicians (n=468, 66.4%) and lived in the United States (n=561, 79.6%). The characteristics of the 705 health care workers are presented in Table 1. Of the 914 comments, 583 (63.8%) were posted from the year 2013 up to the day the COVID-19 outbreak was declared a pandemic by the World Health Organization on March 11, 2020, and 331 (36.2%) were posted during the pandemic (ie, mid-March 2020-December 2020). The lengths of the comments ranged from 1 word to 630 words (mean=81.9, median=60.5, SD 80.4).

Table 1. Characteristics of health care workers who commented on telehealth.

Characteristics	Individuals' frequency, n (%)
Region (N=705)	
Northern America	561 (79.6)
Asia	37 (5.2)
Europe	23 (3.3)
Africa	19 (2.7)
Central or South America	12 (1.7)
Australia or New Zealand	10 (1.4)
Unknown	43 (6.1)
Occupation/position (N=705)	
Physicians	468 (66.4)
Nurses	89 (12.6)
Health administration professionals	42 (6.0)
Other health occupation	105 (14.9)
Unknown	1 (0.1)
Physicians' medical specialties (N=468)	
Family medicine/general practice	106 (22.7)
Clinical medicine specialties	254 (54.3)
Surgery and surgical specialties	45 (9.6)
Mental health	45 (9.6)
Other specialty	5 (1.1)

Sentiment Scores of Comments

The mixed-effect model with the comment sentiment score as the dependent variable showed that, in general, comments made by health care workers living in the United States were less positive compared to their counterparts residing outside the United States. There were no statistically significant differences in average sentiment scores among occupations (physician, nurse, health administration professional, other) or across years in which comments were posted (2013-2020). Hence, overall, the sentiment scores of comments posted during the pandemic were comparable to those of comments made before the pandemic. See Table 2, Textbox 1, and Figure 2.

Table 2. Fixed-effect estimates of independent variables on sentiment scores of comments about telehealth services.

Independent variable	Coefficient estimate (95% CI)	*P* value
Residence: non-US vs US	0.06 (0.02-0.10)	.003[a]
Occupation: nurse vs physician	0.03 (–0.02-0.08)	.23
Occupation: health administration vs physician	0.03 (–0.04-0.10)	.38
Occupation: other vs physician	0 (–0.05-0.04)	.85
Year 2013 vs 2020	–0.06 (–0.18-0.06)	.31
Year 2014 vs 2020	0.03 (–0.01-0.07)	.21
Year 2015 vs 2020	–0.01 (–0.06-0.04)	.62
Year 2016 vs 2020	0.01 (–0.03-0.06)	.58
Year 2017 vs 2020	0.01 (–0.06-0.08)	.81
Year 2018 vs 2020	0.04 (–0.13-0.21)	.63
Year 2019 vs 2020	–0.03 (–0.09-0.02)	.28

[a]*P*<.01.

Textbox 1. Model statistics.

Null model likelihood ratio:

χ^2=9.80, *df*=1, *P*=.002

Random effects:

z=2.83, *P*=.002[a]

Figure 2. Average sentiment score per sentence of comments about telehealth by occupation and by residence.

Topics of Discussions

STM identified 13 topics presented in Table 3. The relationships between the prevalence of the identified topics and the characteristics of the comments or commenters are presented in Table 4. There was no correlation among the 13 identified topics, that is, all pairwise correlations among topics were less than 0.15. The 3 most discussed topics were virtual visits (topic 1; 10.2% of comments), which are live encounters between a patient and a health care worker through a telecommunication channel (video, telephone, etc); the potential benefits of telehealth (topic 2; 10%); and criticisms of telehealth (topic 3; 9.4%).

Table 3. Identified topics from comments (N=914) about telehealth made by health care workers.

Topic	Percentage of Topic in Corpus	Top 10 highest-probability words	Quotes
Topic 1: virtual visits	10.2	visit, patient, virtual, office, time, doctor, video, work	• *My physician's office is seeing patients using virtual visits, when possible, through May.* • *Anyone who thinks a virtual office visit provides the same quality of care as an in-person office visit is seriously deluded. Yet, there clearly is a benefit from this technology in terms of accessibility, and I have used it myself to help people who are unable to visit my office because of debility.*
Topic 2: benefits of telehealth	10.0	care, patient, technology, good, provider, great, telemedicine, service	• *Telemedicine is good if a patient unable to go to their doctors because of many reasons.* • *You are missing the point of telemedicine. It is designed to provide quicker service, [lessen the] burden of travel expenses for the patient, and increase the point of access of care to allow observation and improved treatment plan adherence. Overall, if this is done responsibly, more patients are provided better service, [and] compliance recidivism decreases.*
Topic 3: criticisms of telehealth	9.4	telemedicine, patient, exam, physical, care, antibiotic, medicine, physician	• *I would opt for physical examination of the physical body, rather than virtual examination via telemedicine, every time.* • *There are already studies that show that clinicians using telemedicine are more likely to (over)prescribe antibiotics.*
Topic 4: insurance payment for telehealth services	8.8	patient, phone, care, time, pay, make, insurance, practice	• *I've found 90% of cases can be managed over [the] phone or videoconference. If they offer similar reimbursement from insurance, this may be the way to practice in the future.* • *The government allows doctors to practice telemedicine across state lines but won't allow insurance companies to bid and provide across state lines?*
Topic 5: telehealth over the phone	8.0	telehealth, physician, health, work, patient, call, medicine, clinic	• *Vomiting and diarrhea on Saturday morning. Call primary care physician (35-year history as doctor/patient) for prescription. Answering service: "Sorry, can't prescribe by phone. Go to convenient care." "Can't. Too sick." (Silence.) "Hello, telehealth."* • *We do a follow-up telehealth call to reach out to review labs, etc, with the patients about 2 weeks later. Great for our shut-in patients. They are willing to pay the extra fees associated for the convenience of this service.*
Topic 6: legal dispositions for telehealth	7.6	state, medicine, license, practice, board, health care, patient, telemedicine	• *The current system requires that doctors who practice telemedicine need a license in each state where potential telemedicine patients are. Consequently, a telemedicine doctor in Utah [who] treats patients in Wyoming needs both state licenses.* • *Can a physician licensed in a foreign country provide telehealth services to patients in said country while being on US soil?*
Topic 7: practice of medicine in the era of telehealth	7.4	medicine, patient, doctor, clinical, care, practice, physician, service	• *Let the primary physician do the family practice and refer the cases to the respective specialties. Otherwise, patients will become "Amish."* • *The inhibition of the nonavailability of direct physical assessments is the major legal hurdle for legalizing the authority of prescribing medicines to telemedicine physicians. But dear, this needs to be seen as follows: To overcome these inhibitions, a doctor who is competent in subject could assess the complaints, in the absence of physical findings, by virtue of [their] experience in assessing the clinical value of complaints of telepatients, who currently use teleservices for [a] second opinion.*
Topic 8: moving consultations to the patients' home	7.4	people, patient, health care, work, year, home, telemedicine, health	• *Allowing the patient to be seen in the familiar comfort of their home seems to provide a more optimal environment for cooperation.* • *Huge opportunity for technology to improve health care delivery. The fact that 90% of people have a smartphone opens up endless possibilities to improve how we interact with and support patients. Despite the long-standing emphasis on traditional office visits, the effectiveness of these encounters is not reassuring.*

Topic	Percentage of Topic in Corpus	Top 10 highest-probability words	Quotes
Topic 9: impact on patient-physician relationship	7.1	physician, patient, primary care, relationship, medicine, system, health care, doctor	• *This certainly creates distance between the patient and physician, decreasing the interpersonal relationship and the comfort of human touch, which is tremendously important in [a] bedside manner.* • *Believe it or not, there is no perfect technology that will replace the personal relationship of a doctor with the patient.*
Topic 10: physical examination in the era of telehealth	6.4	patient, exam, care, hand, medicine, system, test, doctor	• *The basic tenet of medicine is laying your hands on the patient and examining what's wrong. Telemedicine is junk medicine except for maybe psychiatric consultations. It will drive up costs gradually, with doctors ordering unnecessary diagnostic tests and reviewing results online and treating results instead of the patient. Instead of unifying care and the concept of a medical home, it will splinter care.* • *In-office visits will soon become archaic remnants of a past rooted in blindly following the old ways. Why would a patient be subject to the time, inconvenience, and danger by subjecting themselves to an in-office visit? And likewise, why would the health care system subject itself to the expense of maintaining staff and facilities necessary for in-office visits? Use of facilities should be restricted to hands-on procedures and assessments only or those requiring specialized equipment.*
Topic 11: impact of telehealth on care quality	6.2	doctor, patient, time, medicine, bad, family, day, registered nurse (RN)	• *The issue of bad medicine is rampant; a consumer requests something, and the prescribing doctor writes for it to make them happy. Telemedicine just makes this problem worse.* • *Medical boards will become involved in it at some point when a bad outcome occurs and family members get upset over that bad outcome. There will have to be some sort of informed consent that the patient will have to accept the risks of no physical assessment and perhaps an incorrect diagnosis.*
Topic 12: issues related to telehealth patient appointment and follow-up	5.9	patient, phone, call, appointment, follow, issue, physician, telemedicine	• *It is also burdensome to contact any health care provider via a telephone call. So, what are patients and health care providers without access to the necessary technology supposed to do regarding annual follow-up appointments and other medical visits requiring actual physical re-examinations.* • *Physicians should certainly be reimbursed for patient care, including videoconferencing, and follow-up phone calls.*
Topic 13: advantages of telehealth for remote and rural areas	5.6	telemedicine, patient, rural, clinic, medicine, hospital, care, remote	• *Definitely virtual doctors are a boon to remote, rural areas where [a] doctor is a rare commodity.* • *This works great for people living in rural areas, with limited access to health care services.*

Table 4. Coefficient estimates of independent variables across the 13 identified topics.

Independent variables	Topic 1: virtual visits, β; *P* value	Topic 2: benefits of telehealth, β; *P* value	Topic 3: criticisms of telehealth, β; *P* value	Topic 4: payment of services, β; *P* value	Topic 5: telehealth over the phone, β; *P* value	Topic 6: legal dispositions for telehealth, β; *P* value	Topic 7: practice of medicine, β; *P* value	Topic 8: Moving consultations into patients' home, β; *P* value	Topic 9: impact on patient-physician relationship, β; *P* value	Topic 10: impact on physical examination, β; *P* value	Topic 11: impact on quality of care, β; *P* value	Topic 12: patient appointment and follow-up, β; *P* value	Topic 13: telehealth for remote and rural areas, β; *P* value
During vs before pandemic	0.05; .002[a]	–0.03; .04[b]	0.10; .04[b]	0.02; .31	0.04; .01[a]	–0.01; .45	–0.06; <.001[a]	0.07; <.001[a]	–0.02; .09	–0.04; .02[b]	–0.04; .01[a]	0.01; .48	0.08; <.001[a]
Sentiment score	0.04; .32	0.35; <.001[a]	–0.14; <.001[a]	0; .97	–0.03; .38	0; .98	0.05; .25	0.04; .25	–0.08; .03[b]	–0.06; .12	–0.12; <.001[a]	–0.03; .19	–0.02; .56
Pandemic[b] sentiment score	0.06; .38	–0.19; .004[a]	0.05; .42	0.05; .48	–0.05; .43	0; .97	0; .99	0.02; .79	–0.02; .80	–0.03; .61	0.12; .04[b]	–0.01; .89	0; .98
Residence: non-US vs US	–0.02; .29	0.01; .46	–0.03; .19	–0.06; .004[a]	–0.01; .63	–0.06; .002[a]	0.15; <.001[a]	0.02; .28	–0.01; .59	–0.04; .04[b]	–0.01; .63	0.02; .19	0.01; .67
Nurse vs physician	–0.02; .22	0.03; .09	–0.06; .002[a]	–0.04; .05	0.03; .08	–0.03; .20	–0.04; .05	0.04; .05	–0.01; .47	0.04; .05	0; .80	0.05; .004[a]	0; .84
Health administrator vs physician	–0.01; .71	0.03; .25	–0.06; .03[b]	0; .88	–0.02; .49	–0.03; .31	0.02; .62	0.04; .16	–0.02; .46	–0.03; .31	–0.03; .31	0.08; <.001[a]	0.02; .41
Other vs physician	0.03; .09	0.04; .04[b]	–0.01; .54	0; .98	0; .85	–0.02; .24	0.06; .01[a]	0.01; .70	–0.04; .03[b]	–0.04; .04[b]	0; .99	0; .86	–0.03; .11

[a]*P*<.01.

[b]*P*<.05.

The frequently mentioned benefits (topic 2) included the possibility telehealth offers to reach patients who cannot come to the health care facility for some reasons (living with a disability, residing in a remote area, etc), the decrease in the transportation time and cost for the patient, and the improvement in patients' access to care and to medical specialists. Some of these benefits were emphasized during the pandemic:

> *[The] majority of my patients have been so glad that we were able to continue with therapy during the pandemic. Some will be happy to get back to face-to-face therapy, but most have seen no change in the sessions.*

Although the topic on the benefits of telehealth (topic 2) was associated with a positive sentiment score overall, the sentiment score for that topic was less positive during the pandemic compared to its prepandemic value. A deeper examination of the comments revealed that this decrease in the sentiment score may have stemmed from the changes in routine care processes due to pandemic lockdowns and from the technical difficulties encountered by health care workers who had acknowledged the value of telehealth services:

> *Change is difficult, but telehealth works great for the provider as well as the patient.*
>
> *Virtual visits for pediatrics will be quite difficult. Limited virtual consultation may be possible, particularly with visually obvious abnormalities that can be followed visually. I cannot visualize evaluations of children with epilepsy, asthma, autism spectrum, cerebral palsy . . . being done by virtual video visits.*
>
> *Sure, telemedicine during a global pandemic to enable social distancing is to be expected, but it is far from a panacea . . . No one also talks about technological limitations . . . I didn't go into health care to troubleshoot people's internet connections. I'd say 50% of my telemedicine encounters have some sort of technical issue that we spend easily the first 10 minutes dealing with.*

Criticisms of telehealth (topic 3) included the impracticality to perform a physical examination of the patient, which may lead to misdiagnoses, and the overprescription of certain medications by health care providers who used telemedicine (eg, antibiotics).

Health care workers who raised these criticisms opined that telehealth devalues "the art and science" of medicine:

> *Telemedicine makes a mockery of the art of medicine.*
>
> *Say goodbye to the thorough, hands-on assessment and the art of medicine.*

Some commenters argued that the resistance to telehealth from some health care professionals may be due the noninclusion of telehealth services into the curricula of training institutions:

> *The problem with telemedicine is that it flies in the face of how many of us were trained. We were constantly told to examine the patient and pay less attention to labs and radiology.*
>
> *Worst of all, very few have really been trained on telemedicine, and how to handle its limitations and pitfalls.*

Other workers stated that adequate training may help decrease the observed resistance to telehealth services and lead to improved quality of these services:

> *With proper training and some careful forethought, telehealth can be done well for those with serious illness.*
>
> *With appropriate training regarding the manner in which this kind of therapeutic contact is to be carried out, videoconferencing can be a valuable conduit for the delivery of clinical treatment or educational services.*

Other topics raised by commenters included the payment of telehealth services by insurance companies (topic 4; 8.8%) and the legal dispositions necessary for the practice of telehealth (topic 6; 7.6%), especially an interstate medical licence for physicians who practice in the United States. Comments on the lack of proper legal dispositions for telehealth were not specific to health care workers from the United States, however. For instance, a physician from Brazil noted that

> *Here where I am (Brazil), there are a series of legal requirements that still discourage (or scare) a professional using video calls. Psychologists are more advanced on this topic, but doctors are not yet. The COVID-19 pandemic is accelerating the discussion, but there are still many barriers and doubts.*

Virtual visits (topic 1), criticisms of telehealth (topic 3), telehealth services (topic 5), moving consultations to patients' homes (topic 8), and the advantages of telehealth for remote areas (topic 13) were more frequently discussed during the COVID-19 pandemic compared to the prepandemic period. In contrast, the benefits of telehealth (topic 2), the practice of medicine (topic 7), physical examination in the era of telehealth (topic 10), and the impact of telehealth on care quality (topic 11) were less frequently discussed during the pandemic compared to before the pandemic. Compared to their counterparts living in other countries, health care workers in the United States discussed insurance payment of telehealth services (topic 4), the legal dispositions for telehealth (topic 6), and physical examination in the era of telehealth (topic 10) more frequently. Non-US health care workers, instead, discussed the practice of medicine in the era of telehealth (topic 7) more frequently than their US counterparts. In general, criticisms of telehealth (topic 2) were less frequently discussed by nurses and health administrators compared to physicians. However, nurses and health administrators discussed issues related to telehealth patient appointments and follow-ups (topic 12) more than physicians.

Although the topic of the impact of telehealth on care quality (topic 11) was associated with a negative sentiment score overall, the score was less negative during the pandemic compared to the prepandemic period, signaling a shift in opinion. Such shift can be seen in comments such as the following posted during the pandemic:

> *Having previously been against telemedicine, now I see a lot of its positive qualities, specifically for the elderly and for the young, however, too.*
>
> *The telemedicine services were underutilized. The new pandemic shows the importance of telemedicine. The wider use of smartphones is also helpful in this direction.*

However, the adoption of telehealth services by some health care workers may not outlast the pandemic, as illustrated by this statement posted by a physician:

> *The telemedicine visit was important for the first 3 months of the pandemic. However, as soon as we were able to see patients in the office, then that became my preferred method to evaluate patients. I agree that [a] lack of the physical exam was the major reason to encourage in-office visits, even with stable, chronic disease patients.*

Discussion

Principal Findings

The objective of this research was to examine the opinions of health care workers about telehealth services and to explore the impact the COVID-19 pandemic may have had on these opinions. The analysis showed that the COVID-19 pandemic did not significantly alter the opinions about telehealth services expressed by health care workers on the discussion forum of the study. Although some topics were more (eg, topic 1: virtual visits) or less (eg, topic 7: practice of medicine) frequently discussed during the pandemic compared to the prepandemic period, the same issues about telehealth services were raised during both time frames. Furthermore, the topic of telehealth benefits (topic 2) was associated with a less positive sentiment during the pandemic compared to before the pandemic. However, the topic of the quality of care of telehealth services (topic 11) had a higher positive sentiment score during the pandemic compared to the prepandemic time frame. This finding may be reflecting the opinion among some health care workers that telehealth can enhance the quality of care only in unusual situations when an in-person encounter between patient and clinician is not feasible, as was the case during the pandemic.

Differences in Comment Sentiment Scores

The sentiment scores of comments posted by health care professionals living in the United States were less positive than those of opinions expressed by their counterparts living in other countries. The lack of additional information about the characteristics of the commenters (eg, sociodemographics, computer knowledge) did not allow the assessment of whether this association between location of residence and sentiment scores was due to differences in sociodemographics or in the knowledge of information technology (IT) between commenters from the United States and their counterparts from other countries. Indeed, previous studies have reported that the attitude toward technology may differ by age, occupational seniority, and level of training in IT [13,14]. It is worth emphasizing that the general trends in sentiment scores do not necessary reflect sentiments about telehealth per se, as demonstrated by the multiple topics and the associated sentiment scores uncovered by the analysis. Rather, these general trends show the aggregate sentiment of the concerns or expectations elicited by discussions of telehealth.

Discussion Topics Associated With Negative Sentiment Scores

Although the pandemic may have offered the opportunity to some health care workers to change their opinions about telehealth services, criticisms of these services were still relatively prevalent among the commenters during the pandemic. Reports on the drawbacks of telehealth, such as the overprescription of antibiotics in direct-to-consumer (DTC) telehealth visits, were frequently mentioned by health care workers who had an unfavorable opinion of telehealth services [15]. DTC telehealth providers offer virtual consultations with health care workers to customers or patients through applications installed on the IT devices of these patients. Patients enter their symptoms or motives of consultation in the application and are then assigned to a health care worker in the network of the DTC telehealth company. The interaction between the patient and the health professional is entirely virtual, and the latter may prescribe a treatment at the end of the encounter. DTC telehealth has often been associated with reduced care quality [16]. It should be mentioned that other types of telehealth services that involve some degree of clinical or paraclinical examination have yielded good quality care [17-19]. Another disadvantage of telehealth that was frequently mentioned before and during the pandemic was the inability to perform a physical examination. Conducting physical examinations has been reported to be perceived by health care workers as part of their identity [20]. Hence, telehealth platforms that do not offer the possibility to perform a physical examination are perceived to threaten that identity. Statements such as “telemedicine makes a mockery of the art of medicine” can be interpreted through the lens of the perceived threat to this occupational identity. It is worth noting that some solutions exist to address the issue of physical exam during telehealth sessions. These solutions include facilitated virtual visits, which require the patient to be examined in a designated facility (originating site), usually by a health care professional (facilitator), during the virtual visit. Information about the physical exam is then transmitted by the facilitator to the patient health provider participating in this telehealth session. Another solution for addressing the issue of physical examination is the use of diagnostic tools (eg, digital stethoscope) that can transmit data from the physical examination remotely [21].

Promoting the Adoption of Telehealth Services

Researchers have shown that the technology acceptance model [22], which posits that the individuals’ perceptions of the technology’s usefulness and ease of use influence their acceptance of that technology, was an adequate framework for predicting health care professionals’ intention to use telehealth services [23,24]. Hence, promoting the usefulness of telehealth services and implementing systems that are perceived as easy to use can boost the adoption of telehealth by health professionals. In this study, topics related to the ease of use of telehealth included not only its limitation for conducting a physical examination (topic 10) but also the lack of appropriate legal and service reimbursement dispositions for its adoption (topics 4 and 6). Hence, enhancing telehealth adoption requires addressing these barriers to its usage. Topics related to the perceived usefulness of telehealth included discussions of its benefits (topic 2) and the advantages it offers for the care of patients living in remote or rural areas (topic 13). However, the promotion of the usefulness of telehealth will need to address the perceived drawbacks of its adoption, namely its reported or perceived negative impact on the quality of care (topic 11) and the practice of medicine (topic 3).

Serrano and Karahanna [25] proposed that a successful session of a telehealth visit requires 3 ingredients: the ability of the patient to relay the relevant information through the technology, the competence of the clinician in eliciting from the patient the appropriate information needed to perform the clinical service, and the capacity of the technology to transmit and present the relevant information to both the clinician and the patient. Health professionals’ frustrations from telehealth may come from the patients’ lack of knowledge on how to use the technology to convey information relevant to their condition. A physician wrote,

> *I put myself in the place of a patient who only knows Skype and who, from one moment to the next, to talk to his doctor, will need to use a platform he has never seen . . . it must be very embarrassing.*

Hence, educating patients on how to use a telehealth platform should be among the measures implemented to improve the quality of telehealth services. Health care professionals providing telehealth services should be trained as well, not only on how to navigate the technological features of the platform, but also on how to foster a virtual environment that promotes a good patient-clinician relationship that can help obtain relevant clinical information from the patient. As demonstrated by Serrano and Karahanna [25], a clinician task-specific capability (eg, competency in patient history enquiry) can compensate for the limitations of technology in performing certain tasks. In addition, the technology should fit the clinical task to be performed. Many of the criticisms against telehealth from clinicians included the difficulty or impossibility to obtain certain information through the platform. Under the task-technology-fit framework, the platform used for the

telehealth should include features that enhance image or sound quality so that the virtual patient-clinician encounter can be as close as possible to an in-person interaction [25]. Additional technological requirements include a secure and adequate internet connection and easy-to-use software [21]. There are platforms that allow performing a certain number of physical examinations (called provider access software), but their capabilities are still limited [21]. Finally, a technological support team should be available when the system malfunctions.

Factors that may improve health care workers' perceptions about telehealth include involving them in the design and implementation of the organization's telehealth platforms [26], having opinion leaders express their support for these services and making resources and technical support available [27], and training adequately workers in utilizing these services [28].

The absence of adequate legal dispositions and the differences in telehealth service payment policies by health insurance companies, as noted by some commenters, can hamper the adoption of telehealth [29]. Adler-Milstein et al [30] reported that in the United States, the promotion of private payer reimbursement policies for telehealth services enhances telehealth adoption, while the requirement of an out-of-state licensure for telehealth decreases its adoption. Hence, wider structural changes beyond individual health care organizations are needed as well to foster the adoption of telehealth services.

Limitations

The study had some limitations. First, the sample of health care workers who posted their comments on telehealth are not necessarily representative of the population of health care professionals working in the United States or in other countries. In addition, the relatively small number of comments collected may not have been exhaustive enough to identify other prevalent viewpoints about telehealth. Finally, there is no available method for determining the exact number of topics to be extracted from a corpus. Methods available are all approximate and use different criteria for selecting the number of topics to extract. Thus, it is possible that methods different from the ones used in this paper (semantic coherence, held-out likelihood, and exclusivity) may have yielded a different value for the number of topics to be extracted. Despite these limitations, the analysis of the collected comments provided a valuable insight into the opinions of health care workers about telehealth services over a period of 8 years (2013-2020).

Conclusion

The COVID-19 pandemic and the lockdowns mandated by public officials to control the spread of infection have fostered an interest in expanding telehealth services among health care organizations. However, hurdles to the widespread adoption of these services still remain, including some health care workers' resistance to telehealth, insufficient or inadequate legal dispositions for providing these services, and a lack of coverage of these services by insurance companies. Promoting changes in the unfavorable attitude toward telehealth among health care professionals may require the promotion of evidence-based literature that demonstrates high satisfaction of both patients and clinicians about telehealth. Furthermore, including telehealth service training into the curricula of health care professional training institutions can help prepare these workers perform these services.

Acknowledgments

PN collected the data, conducted the analyses, and wrote the manuscript.

Conflicts of Interest

None declared.

References

1. World Health Organization. Telemedicine: Opportunities and Developments in Member States. Report on the Second Global Survey on eHealth 2010. Geneva: World Health Organization; 2010.
2. Barnett ML, Ray KN, Souza J, Mehrotra A. Trends in telemedicine use in a large commercially insured population, 2005-2017. JAMA 2018 Nov 27;320(20):2147-2149 [FREE Full text] [doi: 10.1001/jama.2018.12354] [Medline: 30480716]
3. Kane CK, Gillis K. The use of telemedicine by physicians: still the exception rather than the rule. Health Aff (Millwood) 2018 Dec;37(12):1923-1930. [doi: 10.1377/hlthaff.2018.05077] [Medline: 30633670]
4. Morgeson FP, Mitchell TR, Liu D. Event system theory: an event-oriented approach to the organizational sciences. AMR 2015 Oct;40(4):515-537. [doi: 10.5465/amr.2012.0099]
5. Zellmer-Bruhn ME. Interruptive events and team knowledge acquisition. Manag Sci 2003 Apr;49(4):514-528. [doi: 10.1287/mnsc.49.4.514.14423]
6. Morgeson FP. The external leadership of self-managing teams: intervening in the context of novel and disruptive events. J Appl Psychol 2005 May;90(3):497-508. [doi: 10.1037/0021-9010.90.3.497] [Medline: 15910145]
7. Hammar Å. Automatic information processing. In: Seel NM, editor. Encyclopedia of the Sciences of Learning. Boston, MA: Springer; 2012:393-394.
8. Office of the National Coordinator for Health Information Technology. Telemedicine and Telehealth. URL: https://www.healthit.gov/topic/health-it-health-care-settings/telemedicine-and-telehealth [accessed 2021-04-06]
9. Rinker T. sentimentr: Calculate Text Polarity Sentiment version 2.7.1. URL: http://github.com/trinker/sentimentr [accessed 2021-04-06]

10. Roberts ME, Stewart BM, Tingley D, Lucas C, Leder-Luis J, Gadarian SK, et al. Structural topic models for open‐ended survey responses. Am J Pol Sci 2014 Mar 06;58(4):1064-1082. [doi: 10.1111/ajps.12103]
11. World Health Organization. WHO Director-General's opening remarks at the Media Briefing on COVID-19: 11 March 2020. URL: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020 [accessed 2021-04-04]
12. Roberts ME, Stewart BM, Tingley D. Stm: An R package for structural topic models. J Stat Softw 2019;91(2):1-40. [doi: 10.18637/jss.v091.i02]
13. Meier ST, Lambert ME. Psychometric properties and correlates of three computer aversion scales. Behav Res Methods Instrum Comput 1991 Mar;23(1):9-15. [doi: 10.3758/bf03203329]
14. Brumini G, Ković I, Zombori D, Lulić I, Bilic-Zulle L, Petrovecki M. Comparisons of physicians' and nurses' attitudes towards computers. Stud Health Technol Inform 2005;116:608-613. [Medline: 16160325]
15. Ray KN, Shi Z, Gidengil CA, Poon SJ, Uscher-Pines L, Mehrotra A. Antibiotic prescribing during pediatric direct-to-consumer telemedicine visits. Pediatrics 2019 May;143(5):e20182491 [FREE Full text] [doi: 10.1542/peds.2018-2491] [Medline: 30962253]
16. Uscher-Pines L, Mulcahy A, Cowling D, Hunter G, Burns R, Mehrotra A. Access and quality of care in direct-to-consumer telemedicine. Telemed J E Health 2016 Apr;22(4):282-287 [FREE Full text] [doi: 10.1089/tmj.2015.0079] [Medline: 26488151]
17. Fang JL, Campbell MS, Weaver AL, Mara KC, Schuning VS, Carey WA, et al. The impact of telemedicine on the quality of newborn resuscitation: a retrospective study. Resuscitation 2018 Apr;125:48-55. [doi: 10.1016/j.resuscitation.2018.01.045] [Medline: 29408329]
18. Chongmelaxme B, Lee S, Dhippayom T, Saokaew S, Chaiyakunapruk N, Dilokthornsakul P. The effects of telemedicine on asthma control and patients' quality of life in adults: a systematic review and meta-analysis. J Allergy Clin Immunol Pract 2019 Jan;7(1):199-216.e11. [doi: 10.1016/j.jaip.2018.07.015] [Medline: 30055283]
19. Burgos F, Disdier C, de Santamaria EL, Galdiz B, Roger N, Rivera ML, e-Spir@p Group. Telemedicine enhances quality of forced spirometry in primary care. Eur Respir J 2012 Jun 10;39(6):1313-1318 [FREE Full text] [doi: 10.1183/09031936.00168010] [Medline: 22075488]
20. Kelly MA, Freeman LK, Dornan T. Family physicians' experiences of physical examination. Ann Fam Med 2019 Jul 08;17(4):304-310 [FREE Full text] [doi: 10.1370/afm.2420] [Medline: 31285207]
21. Baker J, Stanley A. Telemedicine technology: a review of services, equipment, and other aspects. Curr Allergy Asthma Rep 2018 Sep 26;18(11):60. [doi: 10.1007/s11882-018-0814-6] [Medline: 30259201]
22. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
23. Huang J. Innovative health care delivery system: a questionnaire survey to evaluate the influence of behavioral factors on individuals' acceptance of telecare. Comput Biol Med 2013 May;43(4):281-286. [doi: 10.1016/j.compbiomed.2012.12.011] [Medline: 23375377]
24. Kissi J, Dai B, Dogbe CS, Banahene J, Ernest O. Predictive factors of physicians' satisfaction with telemedicine services acceptance. Health Informatics J 2020 Sep 19;26(3):1866-1880 [FREE Full text] [doi: 10.1177/1460458219892162] [Medline: 31854222]
25. Serrano C, Karahanna E. The compensatory interaction between user capabilities and technology capabilities in influencing task performance: an empirical assessment in telemedicine consultations. MISQ 2016 Mar 3;40(3):597-621. [doi: 10.25300/misq/2016/40.3.04]
26. Demiris G. Examining health care providers' participation in telemedicine system design and implementation. 2006 Presented at: American Medical Informatics Association Annual Symposium Proceedings; 2006; Washington, DC p. 906.
27. Ariens LF, Schussler-Raymakers FM, Frima C, Flinterman A, Hamminga E, Arents BW, et al. Barriers and facilitators to eHealth use in daily practice: perspectives of patients and professionals in dermatology. J Med Internet Res 2017 Sep 05;19(9):e300 [FREE Full text] [doi: 10.2196/jmir.7512] [Medline: 28874336]
28. LeRouge C, Garfield MJ, Collins RW. Telemedicine: technology mediated service relationship, encounter, or something else? Int J Med Inform 2012 Sep;81(9):622-636. [doi: 10.1016/j.ijmedinf.2012.04.001] [Medline: 22579395]
29. Dorsey ER, Topol EJ. State of telehealth. N Engl J Med 2016 Jul 14;375(2):154-161. [doi: 10.1056/nejmra1601705]
30. Adler-Milstein J, Kvedar J, Bates DW. Telehealth among US hospitals: several factors, including state reimbursement and licensure policies, influence adoption. Health Aff (Millwood) 2014 Feb;33(2):207-215. [doi: 10.1377/hlthaff.2013.1054] [Medline: 24493762]

Abbreviations

DTC: direct to consumer
IT: information technology
STM: structural topic modeling

Edited by C Basch; submitted 09.04.21; peer-reviewed by X Zhou, V Strotbaum; comments to author 31.05.21; revised version received 11.07.21; accepted 12.07.21; published 25.02.22.

Please cite as:
Nitiema P
Telehealth Before and During the COVID-19 Pandemic: Analysis of Health Care Workers' Opinions
J Med Internet Res 2022;24(2):e29519
URL: https://www.jmir.org/2022/2/e29519
doi:10.2196/29519
PMID:34978532

Original Paper

Monitoring COVID-19 on Social Media: Development of an End-to-End Natural Language Processing Pipeline Using a Novel Triage and Diagnosis Approach

Abul Hasan[1], MSc; Mark Levene[1], PhD; David Weston[1], PhD; Renate Fromson[2], MBBS; Nicolas Koslover[2], BMBCh; Tamara Levene[2], BMBCh

[1]Department of Computer Science and Information Systems, Birkbeck, University of London, London, United Kingdom

[2]Barnet General Hospital, London, United Kingdom

Corresponding Author:
Abul Hasan, MSc
Department of Computer Science and Information Systems
Birkbeck, University of London
Malet Street, Bloomsbury
London, WC1E 7HX
United Kingdom
Phone: 44 020 7631 8147
Email: abulhasan@dcs.bbk.ac.uk

Abstract

Background: The COVID-19 pandemic has created a pressing need for integrating information from disparate sources in order to assist decision makers. Social media is important in this respect; however, to make sense of the textual information it provides and be able to automate the processing of large amounts of data, natural language processing methods are needed. Social media posts are often noisy, yet they may provide valuable insights regarding the severity and prevalence of the disease in the population. Here, we adopt a triage and diagnosis approach to analyzing social media posts using machine learning techniques for the purpose of disease detection and surveillance. We thus obtain useful prevalence and incidence statistics to identify disease symptoms and their severities, motivated by public health concerns.

Objective: This study aims to develop an end-to-end natural language processing pipeline for triage and diagnosis of COVID-19 from patient-authored social media posts in order to provide researchers and public health practitioners with additional information on the symptoms, severity, and prevalence of the disease rather than to provide an actionable decision at the individual level.

Methods: The text processing pipeline first extracted COVID-19 symptoms and related concepts, such as severity, duration, negations, and body parts, from patients' posts using conditional random fields. An unsupervised rule-based algorithm was then applied to establish relations between concepts in the next step of the pipeline. The extracted concepts and relations were subsequently used to construct 2 different vector representations of each post. These vectors were separately applied to build support vector machine learning models to triage patients into 3 categories and diagnose them for COVID-19.

Results: We reported macro- and microaveraged F_1 scores in the range of 71%-96% and 61%-87%, respectively, for the triage and diagnosis of COVID-19 when the models were trained on human-labeled data. Our experimental results indicated that similar performance can be achieved when the models are trained using predicted labels from concept extraction and rule-based classifiers, thus yielding end-to-end machine learning. In addition, we highlighted important features uncovered by our diagnostic machine learning models and compared them with the most frequent symptoms revealed in another COVID-19 data set. In particular, we found that the most important features are not always the most frequent ones.

Conclusions: Our preliminary results show that it is possible to automatically triage and diagnose patients for COVID-19 from social media natural language narratives, using a machine learning pipeline in order to provide information on the severity and prevalence of the disease for use within health surveillance systems.

(J Med Internet Res 2022;24(2):e30397) doi:10.2196/30397

KEYWORDS

COVID-19; conditional random fields; disease detection and surveillance; medical social media; natural language processing; severity and prevalence; support vector machines; triage and diagnosis

Introduction

Overview

During the ongoing coronavirus pandemic, hospitals have been continuously at risk of being overwhelmed by the number of people developing serious illness. People in the United Kingdom were advised to stay at home if they had coronavirus symptoms and to seek assistance through the National Health Service (NHS) helpline if they needed to [1]. Consequently, there is an urgent need to develop novel, practical approaches to assist medical staff. A variety of methods have been recently developed that involve *natural language processing* (NLP) techniques; the concerns of these methods range from the level of the individual (see, for example, [2,3]) up to the population level [4,5].

Herein, we take a diagnostic approach and propose an end-to-end NLP pipeline to automatically triage and diagnose COVID-19 cases from patient-authored medical social media posts. The triage may inform decision makers about the severity of COVID-19, and diagnosis could help in gauging the prevalence of infections in the population. Attempting a clinical diagnosis of influenza, or in our case a diagnosis of COVID-19, purely based on the information provided in a social media post is unlikely to be sufficiently accurate to be actionable at an individual level, since the quality of this information will be typically noisy and incomplete. However, it is not necessary to have actionable diagnoses at the individual level in order to identify interesting patterns at the population level, which may be useful within public health surveillance systems. For example, text messages from the microblogging site Twitter were used to identify influenza outbreaks [6]. In addition, Twitter data in conjunction with a US Centers for Disease Control and Prevention (CDC) data set were used to predict the percentage of influenza-like illness in the US population [7].

One of our key concerns is in the production of a high-quality human-labeled data set on which to build our pipeline. Here, we give a brief overview of our pipeline and how we developed our data set. The first step in the pipeline was attained by developing an annotation application that detects and highlights COVID-19-related symptoms with their severity and duration in a social media post, henceforth collectively termed as *concepts*. During the second step, relations between symptoms and other relevant concepts were also automatically identified and annotated. For example, *breathing hurts* is a symptom, which is related to a body part, the *upper chest area*.

One author manually annotated our data with concepts and relations, allowing us to present posts highlighted with identified concepts and relations to 3 experts, along with several questions, as shown in Figure 1. The first question asked the experts to triage a patient into 1 of the following 3 categories: *Stay at home*, *Send to a GP* (where GP stands for general physician), or *Send to a hospital*. The second question asked to diagnose the likelihood of COVID-19 on a Likert scale of 1-5 [8].

The 3 experts are junior doctors working in the United Kingdom who were redeployed to work on COVID-19 wards during the first wave of the pandemic, between March and July 2020. Their roles involved the diagnosis and management of patients with COVID-19, including patients who were particularly unwell and required either noninvasive or invasive ventilation. There were some training sessions organized for doctors working in COVID-19 wards. However, these were only provided toward the end of the first wave, as there was initially little knowledge of the virus and how to treat it. In the hospital, the doctors followed local protocols, which were adjusted as more experience was gained about the virus.

We also asked the doctors to indicate whether the highlighted text presented is sufficient in reaching their decision in order to understand its usefulness when we incorporate it in the annotation interface. The annotations were found to be sufficient in as many as 85% of the posts, on average, as indicated by the doctors' answers to question 3 in Figure 1.

The posts labeled by the doctors were then used to construct 2 types of predictive machine learning model using *support vector machines* (SVMs) [9,10]; see the Step 4: Triage and Diagnosis subsection in the Methods section. The *triage models* use hierarchical binary classifiers, which consider the risk averseness or tolerance of the doctors when making the diagnosis [11]. The *diagnostic models* first calculate the probability of a patient having COVID-19 from doctors' ratings. The probabilities are then used to construct 3 different decision functions for classifying *COVID* and *NO_COVID* classes; these are detailed in the Problem Setting subsection in the Methods section.

We trained the SVM models in 2 different ways: first with ground-truth annotations and second using predictions from the concept and relation extraction step described before. Predictions obtained from the concept extraction step make use of *conditional random fields* (CRFs) [12]; see the Step 1: Concept Extraction subsection in the Methods section for implementation details. Relations are obtained from these predicted concepts using an unsupervised *rule-based* (RB) classifier [13]; see the Step 2: Relation Extraction subsection in the Methods section.

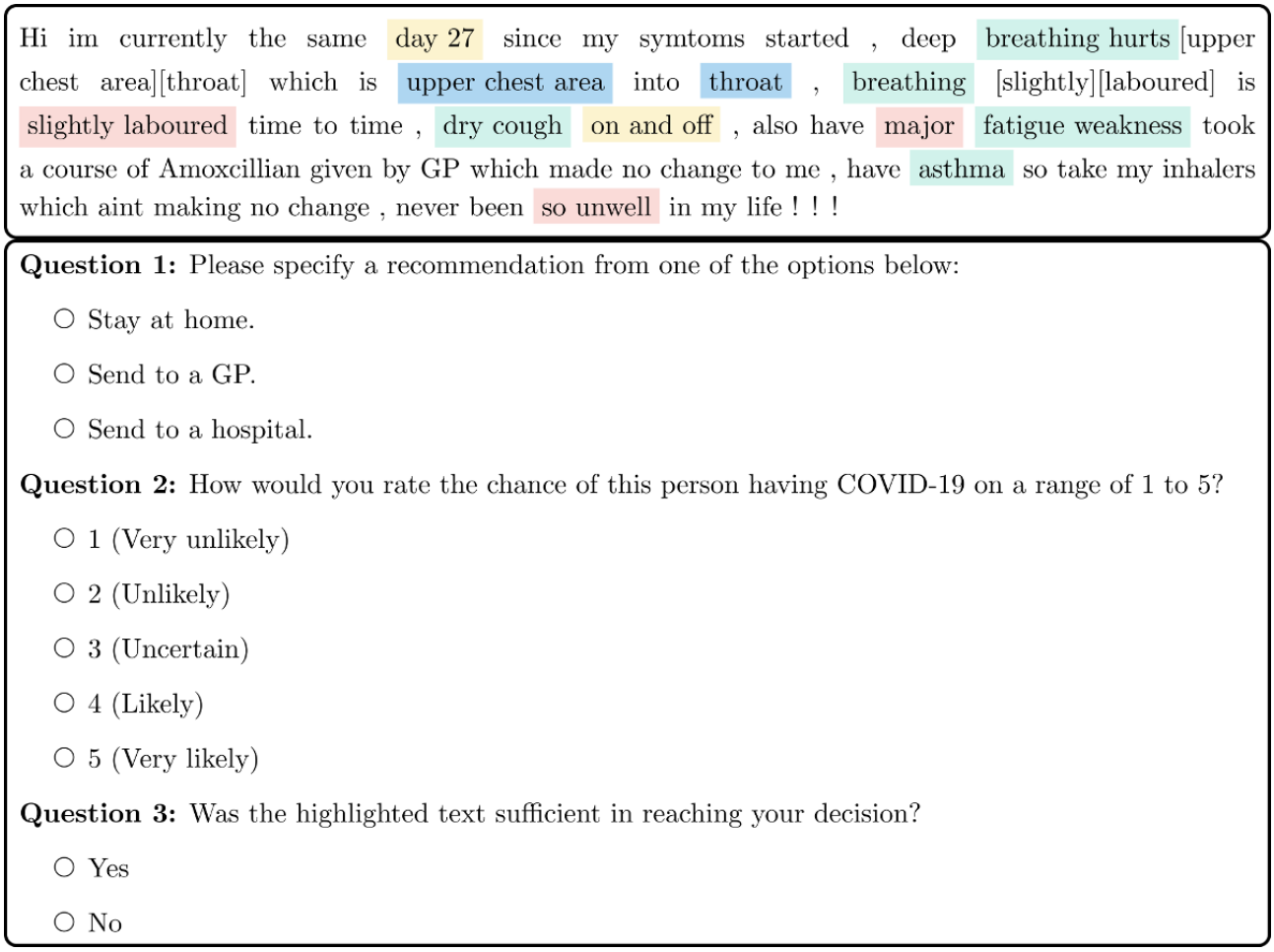
We also discussed the feature importance obtained from the constructed COVID-19 diagnostic models and compared it with the most frequent symptoms from Sarker et al [4] and our data set. We found that symptoms such as anosmia/ageusia (loss of smell/taste) rank in the top 5 most important features, whereas they do not rank in the top 5 most frequent symptoms; see the Discussion section. Overall, we made several contributions as follows:

- We showed that it is possible to take an approach that aims at disease detection to augment public health surveillance systems, by constructing machine learning models to triage and diagnose COVID-19 from patients' natural language narratives. To the best of our knowledge, no other previous

work has attempted to triage or diagnose COVID-19 from social media posts.

- We also built an end-to-end NLP pipeline by making use of automated concept and relation extraction. Our experiments showed that the models built using predictions from concept and relation extraction produce similar results to those built using ground-truth human concept annotation.

Figure 1. A patient-authored social media post is annotated with symptoms (light green), affected body parts (pale blue), duration (light yellow), and severities (pink). The phrases in square brackets show relations between a symptom and a body part/duration/severity when the distance is greater than 1. This annotated post was presented to 3 doctors to triage and diagnose the author of the post by answering questions 1 and 2, respectively. GP: general physician.



Related Work

Data derived from social media have been successfully used to facilitate the detection of influenza epidemics [6,7]. In addition, Edo-Osagie et al [14] provide a thorough review of the use of Twitter in public health surveillance for the purpose of monitoring, detecting, and forecasting influenza-like illnesses. Since the start of the COVID-19 pandemic, a number of mobile application–based, self-reported symptom tools have emerged to track novel symptoms [15]. The mobile application in Menni et al [16] applied logistic regression (LR) to predict the percentage of probable infected cases among the total application users in the United States and United Kingdom combined. Mizrahi et al [17] performed a statistical analysis on primary care electronic health record (EHR) data to find longitudinal dynamics of symptoms prior to and throughout the infection.

At an individual diagnostic level, Zimmerman et al [18] applied classification and regression trees to determine the likelihood of symptom severity of influenza in clinical settings. Moreover, machine learning algorithms, such as decision trees, have shown promising results in detecting COVID-19 from blood test analyses [19]. Here, we focus on features extracted from a textual source to triage and diagnose COVID-19 for the purpose of providing population-level statistics in the context of public health surveillance. Studies related to our work deploy features obtained from online portals, telehealth visits, and structured and unstructured patient/doctor notes from EHRs. In general, COVID-19 clinical prediction models can broadly be categorized into risk, diagnosis, and prognosis models [20].

In Judson et al [21], a portal-based COVID-19 self-triage and self-scheduling tool was used to segment patients into 4 risk categories: emergent, urgent, nonurgent, and self-care, whereas the online telemedicine system in Liu et al [22] used LR to predict low-, moderate-, and high-risk patients by utilizing demographic information, clinical symptoms, blood tests, and computed tomography (CT) scan results.

In Schwab et al [3], various machine learning models were developed to predict patient outcomes from clinical, laboratory,

and demographic features found in EHRs [23]. The authors reported that gradient boosting (XGB), random forests, and SVMs are the best-performing models for predicting COVID-19 test results, hospital admissions, and intensive care unit admissions for positive patients, respectively. A detailed list of clinical and laboratory features can be found in Wang et al [24], where the authors developed predictive models for the inpatient mortality in Wuhan using an ensemble of XGB models. Similarly, in Vaid et al [25], mortality and critical events for patients using XGB classifiers were predicted. Finally, a critical review on various diagnostic and prognostic models of COVID-19 used in clinical settings can be found in Wynants et al [20].

In Wagner et al [26], COVID-19 symptoms from unstructured clinical notes in the EHRs of patients subjected to COVID-19 polymerase chain reaction (PCR) testing were extracted. In addition, COVID-19 SignSym [27] was designed to automatically extract symptoms and related attributes from free text. Furthermore, the study by López-Úbeda et al [28] utilized radiological text reports from lung CT scans to diagnose COVID-19. Similar to our approach, López-Úbeda et al [28] first extracted concepts using a popular medical ontology [29] and then constructed a document representation using word embeddings [30] and concept vectors [28]. However, our methodology differs from theirs with respect to the extraction of relations between concepts, and moreover, our data set, comprising posts obtained from medical social media, is more challenging to work with, since social media posts exhibit greater heterogeneity in language than radiological text reports.

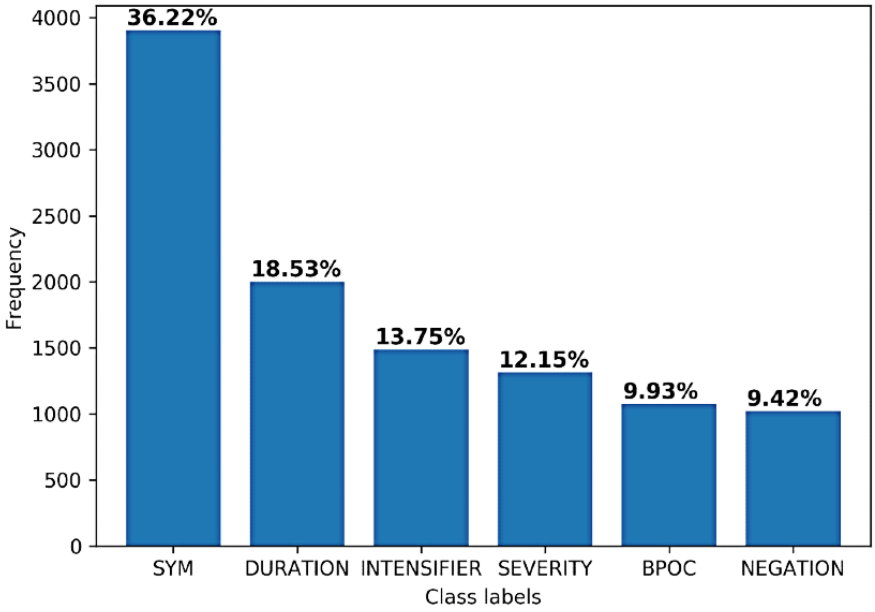
Finally, Sarker et al [4] published a COVID-19 symptom lexicon extracted from Twitter, which we compared our work to in the Discussion section.

Methods

Data

We collected social media posts discussing COVID-19 medical conditions from a forum called *Patient* [31]. This a public forum that was created at the onset of the coronavirus outbreak in the United Kingdom. We obtained permission from the site administrator to scrape publicly available posts dated between April and June 2020. In addition, all user IDs and metadata were removed from the posts for the purpose of the study. After the posts were anonymized, and duplicates were removed, we randomly selected 500 distinct posts. The first author annotated these posts with the classes shown in Figure 2. The class labels represent symptoms and the related concepts: (1) duration; (2) intensifier, which increases the level of symptom severity; (3) severity; (4) negation, which denotes the presence or absence of the symptom or severity; and (5) affected body parts. We also annotated relations between a symptom and other concepts that exist at the sentence level. For example, the relation between a symptom and a severity concept is denoted as *(SYM, SEVERITY)*. The posts were then marked with concepts in different colors, and the relations were placed right after the symptom in square brackets, as shown in Figure 1. Each marked post was presented to the doctors using a web application, and they were asked 3 questions independently; see Figure 1. We called the doctors' answers to questions 1 and 2 as the COVID-19 symptom triage and diagnosis, respectively. Thus, for each post, we had 3 independent answers from 3 doctors, which we denoted as A, B, and C, respectively; these corresponded to the last 3 authors of the paper and were assigned randomly.

Figure 2. Frequency distribution of annotated classes/concepts from the text are shown. We have also shown the percentage of each class after discounting the OTHER labels. The average number of tokens per post was 130.17 (SD 97.83). BPOC: body part, organ, or organ component; SYM: symptoms.



Measurement of Agreement

To measure the agreement between the answers (recommendations and ratings) of the 3 doctors to questions 1 and 2 of Figure 1, we first calculated the proportion of observed agreement (ρ_o), as suggested by de Vet et al [32], who stipulated that Cohen κ is actually a measure of reliability rather than

agreement; we observed that ρ_o was high in all cases, as can be seen in Table 1. We noted that the paradoxical behavior of Cohen κ can arise when the absolute agreement (ρ_o) is high [33]. This may occur when there is a substantial imbalance in the marginal totals of the answers, which we observed in the answers to question 1. Consequently, in addition to Cohen κ, we deployed a common solution to this problem, called the AC1 statistic devised by Gwet and coworkers [34,35].

We found that for question 1, the AC1 measure showed moderate agreement (in the middle of the moderate range) between A and B (0.55) and substantial agreement between A and C (0.72); see Landis and Koch [36] for the benchmark scale for the strength of agreement. For question 2, it turned out that said paradox did not occur, resulting in similar values for κ and AC1. The agreement between A and B (κ=0.64, AC1=0.67) and between B and C (κ=0.64, AC1=0.67) was substantial, while the agreement between A and C (κ=0.40, AC1=0.40) was on the boundary of fair and moderate; see Table 1.

It is important to note that COVID-19 is a novel virus disease, for which the doctors did not have prior experience or training before the first wave of the pandemic, and thus one would expect some difference of opinion. (We bear in mind that in our setting, the doctors can only see the posts and thus cannot interact with the patients as they would in a normal scenario.) Moreover, there are probable differences in risk tolerances between the doctors, which would lead to potentially different decisions and diagnoses.

Table 1. Pairwise agreement between pairs of doctors' answers to questions 1 and 2; see Figure 1 for an example.

Pair	Question 1			Question 2		
	ρ_o	κ	AC1	ρ_o	κ	AC1
AB	0.65	0.26	0.55	0.73	0.64	0.67
BC	0.63	0.14	0.53	0.73	0.64	0.67
AC	0.77	0.28	0.72	0.51	0.40	0.40

Problem Setting

Triage Classification for Question 1

We mapped the doctors' recommendations from question 1 to ordinal values; the options *Stay at home*, *Send to a GP*, or *Send to a hospital* were transformed to the values 1, 2, and 3, respectively. To combine recommendations from 2 or more doctors, we first took their average. This result was rounded to an integer in 1 of 2 ways: either by taking the floor or by taking the ceiling. Considering the risk attitude prevalent among medical practitioners [11], we categorized the ceiling of the average to be *risk averse*, denoted by, for example, AB(R-a), and the floor to be *risk tolerant*, denoted by, for example, AB(R-t). Thus, for each patient's post, we had in total 11 recommendations from 3 doctors for question 1. We constructed a hierarchical classification model for each of these recommendations, where the goal was to classify a post into 1 of the 3 options.

Diagnosis Classification for Question 2

To diagnose whether a patient has COVID-19 from their post, we first estimated the probability of having the disease by normalizing the rating (ie, given a rating, r, the probability of COVID-19, *P*r(COVID|r), which we termed the *ground-truth probability* (GTP), was simply *P*r(COVID|r) = (r – 1)/4.

Given our GTP estimates were discrete, we investigated 3 decision boundaries, denoted by LE, LT, and NEQ, based on a threshold value of 0.5 to classify a post as follows:

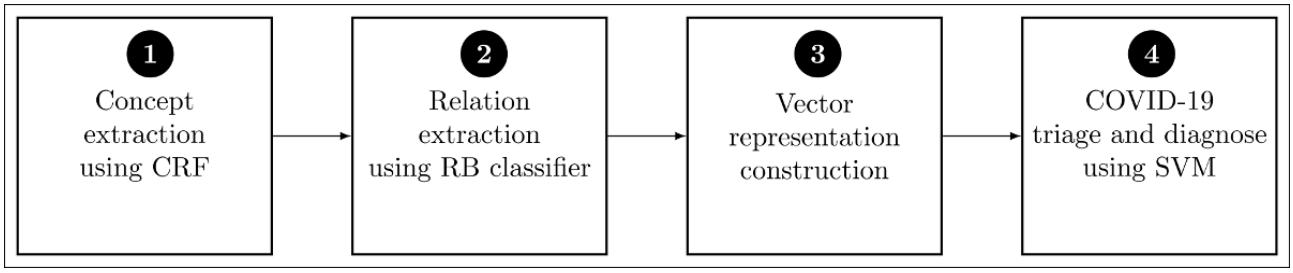
- LE: If Pr(COVID|r)≤0.5, then NO_COVID, else COVID.
- LT: If Pr(COVID|r)<0.5, then NO_COVID, else COVID.
- NEQ: If Pr(COVID|r)<0.5, then NO_COVID, elseif Pr(COVID|r)>0.5, then COVID.

Note that NEQ ignores cases on the 0.5 boundary.

Methodology

A schematic of our methodology to triage and diagnose patients based on their social posts is shown in Figure 3. Here, the circles denote the steps followed in the pipeline. We now detail each of these steps.

Figure 3. A block diagram of the COVID-19 triage-and-diagnosis text processing pipeline. CRF: conditional random field; RB: rule based; SVM: support vector machine.



Step 1: Concept Extraction

In the first step, we preprocessed each patient's post by splitting it into sentences and tokens using General Architecture for Text Engineering (GATE) software's (University of Sheffield) [37] built-in NLP pipeline. For each token in a sentence, we built discrete features that signal whether the token is a member of 1 of the following dictionaries: (1) Symptom, (2) Severity, (3) Duration, (4) Intensifier, and (5) Negation. The dictionaries were built by analyzing the posts while annotating them. We also utilized the MetaMap system [29], assuming that it contains all the necessary technical terms, to map tokens to 3 useful semantic categories: *Sign or Symptom*; *Disease or Syndrome*; and *Body Part, Organ, or Organ Component*. Due to the assumption regarding medical terms, the system does not expect any new additional terms, and thus we were justified in extracting concepts and relations in preprocessing steps. The preprocessed text was then used to build a concept extraction module to recognize the classes, shown in Figure 2, by applying a CRF [12]. A detailed description of our CRF training methodology can be found in Hasan et al [38]. The extracted concepts were then used for our next step to recognize the relations between concepts.

Step 2: Relation Extraction

The semantic relation between a symptom and other concepts, which we formally termed *modifiers*, was resolved using an unsupervised RB classifier algorithm. We first filtered all symptom and modifier pairs from a sentence within a predefined distance and then selected the closest modifier to a symptom to construct a relation. In total, we extracted 5 kinds of relations as follows: *(SYM, SEVERITY)*, *(SYM, DURATION)*, *(SYM, BPOC)*, *(SYM, NEGATION)*, and *(SYM, ?)*—here, SYM and BPOC refer to symptoms, and body part, organ, or organ component, respectively.

The severity modifiers were mapped to a scale of 1-5; the semantic meaning of the scale was *very mild*, *mild*, *moderate*, *severe*, and *very severe*, respectively. The duration modifiers were also mapped to real values in chunks of weeks. So, for example, *10 days* was mapped to the value *1.43*.

Step 3: Vector Representation

Fixed-length vector representations suitable as input for SVM classifiers were built as follows:

- *Symptom-only* vector representation: Let $<S_0, S_1, \ldots, S_n>$ be a vector of symptoms constructed from the symptom vocabulary; for our data set, the number of unique symptom words/phrases was n=871. To construct the vector representation for a post, we extracted the concept, *SYM*, and the relation (*SYM*, *NEGATION*) and set S_i to 1, 0, or –1 according to whether the symptom was present, not present, or negated, respectively.
- *Symptom-modifier relation vector* representation: The symptom-modifier relation vector is a much larger vector than the symptom-only vector and comprises 3 appended vectors containing (1) the absence or presence of 110 unique body parts, (2) the absence or value of a symptom duration, and (3) the absence, negation, or value or a symptom severity.

Step 4: Triage and Diagnosis

We utilized SVM classification and regression models to triage and diagnose patients' posts, respectively, from the vector representations described earlier. For question 1, the recommendation from a doctor or a combination of doctors was the class label of the post; see the Problem setting subsection in the Methods section for a description. To build a binary classifier, we first combined the *Send to a GP* and *Send to a hospital* recommendations to represent a single class, *Send*. The SVM was trained to distinguish between the *Stay at home* and the *Send* options; we called this *SVM classifier 1*. Next, the posts labeled as *Stay at home* were discarded and *SVM classifier 2* was built utilizing the remaining posts to classify the *Send to a GP* and *Send to a hospital* recommendations. This resulted in a hierarchical classifier for COVID-19 triage.

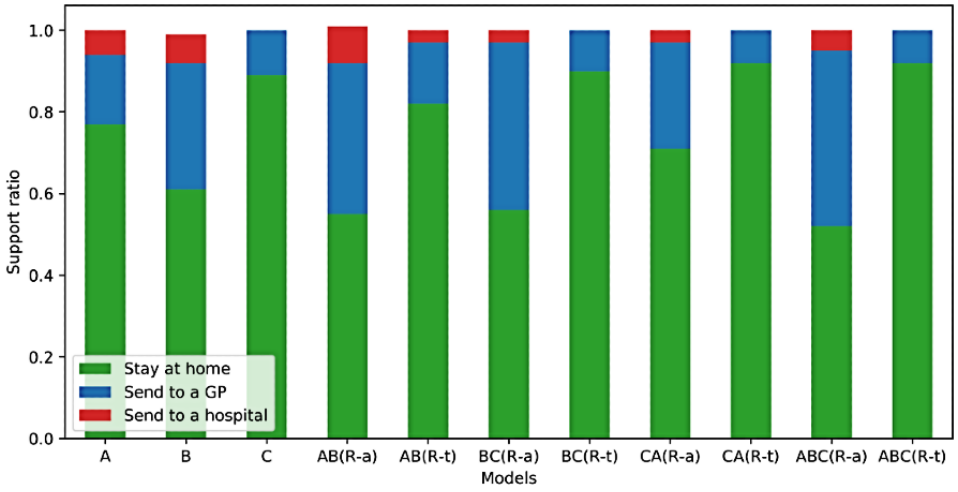
For diagnosing COVID-19 cases, we deployed a variant of the SVM, called *support vector regression* (SVR) [9], to estimate the probability of COVID-19. We used the GTP that was derived from answers to question 2 as the dependent variable. SVR takes as input a high-dimensional feature vector, such as a symptom-only or a symptom-modifier relation vector representation, as described earlier. Classification was performed using the 3 decision functions, LE, LT, and NEQ, described previously.

Results

Evaluation

We evaluated the performance of the CRF and SVM classification algorithms using the standard measures of precision, recall, and macro- and microaveraged F_1 scores [39]. Macroaveraged scores were computed by considering the score independently for each class and then taking the average, while microaveraged scores were computed by considering all the classes together. As our data set was not balanced with *COVID* and *NO_COVID* classes, as can be seen in Figure 4, and we wished to give equal weight to all instances, we reported microaveraged scores for the SVR classification. In contrast, in the case of concept extraction, the *Other* class dominated. So, in this case, we reported the macroaveraged scores for the CRF classification results.

Figure 4. Support ratio of triage classes across models for question 1 classification tasks. Absolute numbers for the "Send to a hospital" class in test sets were as follows: A=10, B=12, AB(R-a)=14, AB(R-t)=5, BC(R-a)=6, AC(R-a)=5, and ABC(R-a)=9; the value for the remaining models was 0. GP: general physician.



Experimental Setup

For the CRF, we reported 3-fold cross-validated macroaveraged results. Specifically, we trained each fold by a Python wrapper [40] for CRFsuite; see Okazaki [41]. For relation extraction, we ran our unsupervised RB algorithm on the 500 posts and calculated the F_1 scores by varying distances considering the 2 cases with and without stop words.

We constructed SVM binary classifiers, SVM classifier 1 and SVM classifier 2, using the Python wrapper for LIBSVM [42] implemented in Sklearn [43] with both linear and Gaussian *radial basis function* (RBF) kernels [10]. Similarly, SVR [44] was implemented using LIBSVM and was built with both linear and RBF kernels. The hyperparameters (C=10 for the penalty, γ=0.01 for the RBF kernel, and ε=0.5 for the threshold) were discovered using a grid search [43].

We simulated 2 cases for COVID-19 triage and diagnosis. First SVM and SVR models were trained with the ground truth to examine the predictive performance when they are deployed as stand-alone applications. Second, when trained with the predictions from the CRF and RB classifier, they resembled an end-to-end NLP application. To obtain a comparable result, the models were always tested with the ground truth. As a measure of performance, we reported macro- and microaveraged F_1 scores for SVM classifiers and SVR, respectively.

Evaluation Outcomes

The concept and relation extraction phases produced excellent and good predictive performances, respectively; see Tables 2 and 3. The triage classification results from question 1 are shown in Tables 4 and 5; the full enumeration can be seen in the first column. When we trained the models with the symptom-modifier vector representations from the ground truth, the results of SVM classifier 1 and SVM classifier 2 were in the range of 72%-93% and 83%-96%, respectively. The symptom-only vector representations produced results in the range of 71%-94% and 79%-95%. These results suggested that we can achieve good predictive performance for classifying *Stay at home* and *Send* and for *Send to a GP* and *Send to a hospital*. In general, risk-tolerant models achieved better performance than risk-averse models. However, since in the test set, posts with the label *Send to a hospital* were missing for some models (as can be seen from Figure 5), we could not report them. We reported macroaveraged F_1 score results since question 1 was framed as a decision problem, where weights for the classes are a priori equal. The results obtained after training with CRF predictions were in similar ranges for both representations and classifiers. This is important, because it indicated that an end-to-end NLP application is likely to produce similar predictive performance.

Regarding question 2, when we trained the models with the symptom-modifier vector representation from the ground truth, the results of COVID-19 diagnosis were in the range of 72%-87%, 61%-76%, and 74%-87% for the LE, LT, and NEQ decision functions, respectively; see Table 6. The symptom-only vector representation produced results in the range of 70%-88%, 59%-79%, and 74%-87% for the LE, LT, and NEQ decision functions, respectively.

In general, NEQ models perform better due to the omission of borderline cases where the GTPs are exactly 0.5. The support ratios for each model for different decision functions are shown in Figure 4. When we trained the models with the symptom-modifier vector representation from the CRF predictions, the results were in the range of 68%-86%, 64%-76%, and 73%-87% for the LE, LT, and NEQ decision functions, respectively. This indicated that for diagnosis as well as triage, an end-to-end NLP application is likely to perform similarly to stand-alone applications. Here, we reported microaveraged F_1 scores since, in our data set, *NO_COVID* cases dominated; this largely resembled the natural distribution in the population, where people who tested positive for coronavirus are a relatively low percentage in the whole population, even when the prevalence of the virus is high.

Finally, we trained our models using a linear kernel but found that the RBF dominates in most of the cases; however, linear kernels are useful in finding feature importance [45].

Table 2. Concept extraction using CRF[a] on 3-fold cross-validation.

Label	Precision	Recall	F_1 score	Support
SYM[b]	0.94	0.97	0.95	1300
SEVERITY	0.80	0.79	0.79	437
BPOC[c]	0.92	0.83	0.87	356
DURATION	0.87	0.91	0.89	667
INTENSIFIER	0.88	0.97	0.92	494
NEGATION	0.83	0.89	0.86	338
OTHER	0.99	0.98	0.98	16892
Macroaverage	0.89	0.89	0.89	__[d]

[a]CRF: conditional random field.

[b]SYM: symptoms.

[c]BPOC: body part, organ, or organ component.

[d]Not applicable.

Table 3. Relation extraction using RB[a] classifier results on 3-fold cross-validation.

Distance	With stop words			Without stop words		
	Precision	Recall	F_1 score	Precision	Recall	F_1 score
2	0.74	0.63	0.68	0.74	0.64	0.69
3	0.75	0.67	0.71	0.75	0.67	0.71
4	0.75	0.69	0.72	0.75	0.69	0.72
5	0.75	0.71	0.73	0.74	0.71	0.73
6	0.74	0.72	0.73	0.74	0.72	0.73
7	0.73	0.73	0.73	0.73	0.73	0.73

[a]RB: rule based.

Table 4. Question 1: hierarchical classification results for the RBF[a] kernel using the symptom-modifier relation vector.

Model	SVM[b] classifier 1			SVM classifier 2		
	Precision	Recall	F_1 score	Precision	Recall	F_1 score
Trained on the ground truth						
A	0.82	0.91	0.86	0.73	0.95	0.83
B	0.73	0.77	0.75	0.81	0.99	0.89
C	0.85	0.98	0.91	__[c]	—	—
AB(R-a)	0.70	0.75	0.72	0.80	0.96	0.88
AB(R-t)	0.84	0.96	0.89	0.85	1.00	0.92
BC(R-a)	0.72	0.75	0.73	0.92	1.00	0.96
BC(R-t)	0.86	0.99	0.92	—	—	—
AC(R-a)	0.79	0.87	0.83	0.89	1.00	0.94
AC(R-t)	0.88	0.98	0.93	—	—	—
ABC(R-a)	0.70	0.76	0.73	0.89	0.99	0.93
ABC(R-t)	0.88	0.99	0.93	—	—	—
Trained on the CRF[d] predictions						
A	0.81	0.89	0.85	0.72	0.91	0.80
B	0.74	0.74	0.74	0.81	0.99	0.89
C	0.85	0.96	0.90	—	—	—
AB(R-a)	0.73	0.71	0.71	0.81	0.96	0.88
AB(R-t)	0.84	0.94	0.88	0.84	1.00	0.92
BC(R-a)	0.74	0.71	0.72	0.92	1.00	0.96
BC(R-t)	0.88	0.98	0.93	—	—	—
AC(R-a)	0.81	0.85	0.83	0.89	1.00	0.94
AC(R-t)	0.88	0.98	0.93	—	—	—
ABC(R-a)	0.72	0.72	0.72	0.89	1.00	0.94
ABC(R-t)	0.89	0.98	0.93	—	—	—

[a]RBF: radial basis function.

[b]SVM: support vector machine.

[c]Not applicable.

[d]CRF: conditional random field.

Table 5. Question 1: hierarchical classification results for the RBF[a] kernel using the symptom-only vector.

Model	SVM[b] classifier 1			SVM classifier 2		
	Precision	Recall	F_1 score	Precision	Recall	F_1 score
Trained on the ground truth						
A	0.83	0.91	0.87	0.74	0.85	0.79
B	0.71	0.81	0.76	0.81	0.98	0.89
C	0.87	0.97	0.92	__[c]	—	—
AB(R-a)	0.69	0.75	0.72	0.83	0.96	0.89
AB(R-t)	0.85	0.94	0.89	0.85	1.00	0.92
BC(R-a)	0.71	0.79	0.75	0.92	0.99	0.95
BC(R-t)	0.88	0.98	0.93	—	—	—
AC(R-a)	0.80	0.86	0.83	0.89	1.00	0.94
AC(R-t)	0.90	0.98	0.94	—	—	—
ABC(R-a)	0.68	0.74	0.71	0.90	1.00	0.95
ABC(R-t)	0.90	0.98	0.94	—	—	—
Trained on the CRF[d] predictions						
A	0.84	0.89	0.87	0.74	0.82	0.78
B	0.74	0.79	0.77	0.82	0.98	0.89
C	0.86	0.95	0.90	—	—	—
AB(R-a)	0.72	0.76	0.73	0.83	0.92	0.87
AB(R-t)	0.87	0.93	0.90	0.84	0.98	0.90
BC(R-a)	0.72	0.78	0.75	0.92	0.99	0.95
BC(R-t)	0.87	0.97	0.92	—	—	—
AC(R-a)	0.80	0.86	0.83	0.89	1.00	0.94
AC(R-t)	0.89	0.95	0.92	—	—	—
ABC(R-a)	0.71	0.76	0.73	0.89	0.99	0.93
ABC(R-t)	0.90	0.95	0.92	—	—	—

[a]RBF: radial basis function.

[b]SVM: support vector machine.

[c]Not applicable.

[d]CRF: conditional random field.

Figure 5. Support ratio of diagnosis classes across models and 3 decision functions for question 2 classification tasks.

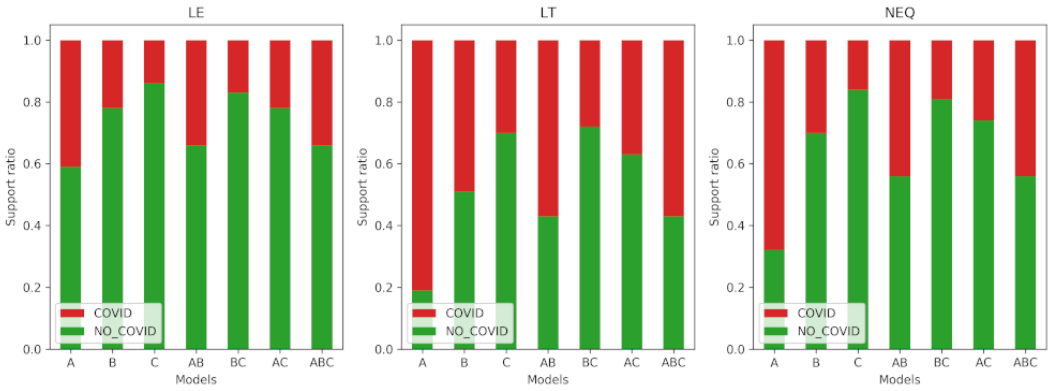


Table 6. Question 2: microaveraged F_1 score results for different models and decision functions. Here, A, B, and C are 3 medical doctors (abbreviated as Dr) who took part in the experiment.

Model	Symptom-modifier vector			Symptom-only vector		
	LE	LT	NEQ	LE	LT	NEQ
Trained on the ground truth						
A	0.72	0.61	0.78	0.70	0.59	0.74
B	0.78	0.61	0.76	0.78	0.62	0.77
C	0.87	0.75	0.87	0.88	0.75	0.87
AB	0.72	0.66	0.74	0.74	0.65	0.75
BC	0.84	0.76	0.84	0.85	0.79	0.86
AC	0.81	0.73	0.81	0.83	0.74	0.83
ABC	0.74	0.67	0.76	0.75	0.67	0.77
Trained on the CRF[a] predictions						
A	0.68	0.64	0.76	0.50	0.79	0.74
B	0.76	0.64	0.77	0.78	0.57	0.74
C	0.86	0.75	0.87	0.87	0.74	0.86
AB	0.70	0.65	0.73	0.71	0.66	0.74
BC	0.83	0.76	0.83	0.85	0.78	0.86
AC	0.80	0.74	0.82	0.80	0.73	0.81
ABC	0.72	0.69	0.76	0.74	0.69	0.77

[a]CRF: conditional random field.

Discussion

Principal Findings

This study demonstrates the potential to triage and diagnose COVID-19 patients from their social media posts. We presented a proof-of-concept system to predict a patient's health state by building machine learning models from their narrative. The models were trained in 2 ways: using (1) ground-truth labels and (2) predictions obtained from the NLP pipeline. Trained models are always tested on ground-truth labels. We obtained good performances in both cases, which indicates that an automated NLP pipeline could be used to triage and diagnose patients from their narrative; see the Evaluation Outcomes subsection in the Results section. In general, health professionals and researchers could deploys triage models to determine the severity of COVID-19 cases in the population and diagnostic models to gauge the prevalence of the pandemic.

Comparison With Prior Work

To quantify the important predictive features in the training set, we experimented with COVID-19 diagnosis using linear kernel SVR regression. More specifically, we used the symptom-only vector representation constructed from the ground truth. We summed feature weights for each S_i in $<S_0, S_1, \ldots, S_n>$ from the 7 models and the 3 decision functions; see the Methods section. The features were then mapped to the categories found in the Twitter COVID-19 lexicon complied by Sarker et al [4]. The top 5 important features in our data set were *cough*, *anosmia/agusia*, *dyspnea*, *pyrexia*, and *fatigue*. Mizrahi et al [17] quoted 4 of these symptoms as the most prevalent coronavirus symptoms, strongly correlating with our findings.

To compare our importance ranking with that of Sarker et al's [4] frequent categories, we compiled the corresponding frequencies of our 5 most important symptoms. Normalized weights and frequencies were then plotted in Figure 6. The top-left stacked bar chart compares our 5 most important features with Sarker et al's [4] frequencies. Cough was the most important symptom from our data set, where it was the second-most frequent. Anosmia/ageusia ranked second in our importance list, while it was seventh in the most frequent list. Pyrexia came first and fourth in both the frequent and importance lists, respectively.

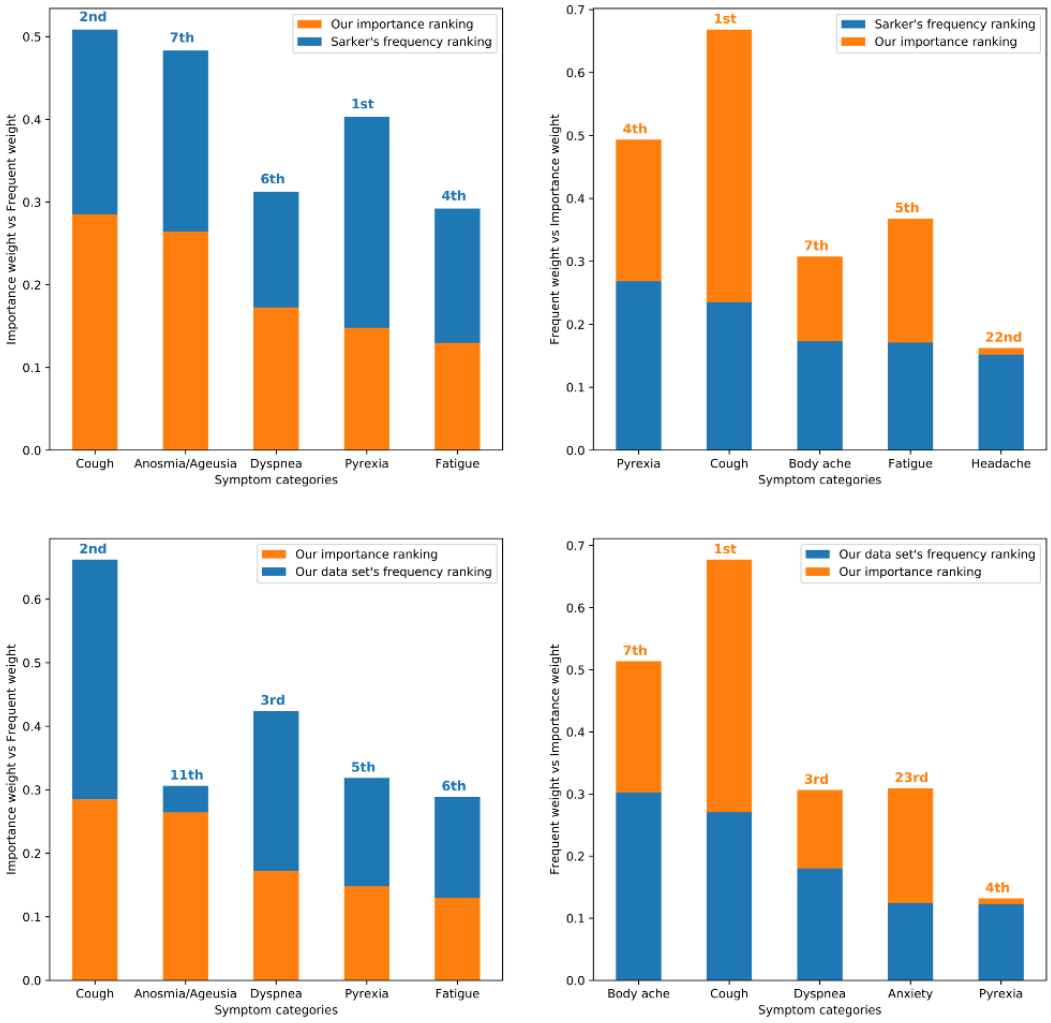
The top-right chart in Figure 6 shows a comparison between Sarker et al's [4] frequency ranking and our importance ranking. Here, we selected the top 5 most frequent symptoms from Sarker et al's [4] frequency list and normalized them. These are *pyrexia*, *cough*, *body ache*, *fatigue*, and *headache*. We took the corresponding importance weights of these symptoms and plotted them in a stacked bar chart. Here, headache ranked 22nd in our importance ranking, while it was 5th in the frequency ranking. We found a large difference between the 2 rankings, implying that the top-most frequent symptoms are not necessarily the most important ones.

Next, we compared our most important feature weights with our data set's frequency ranking using the methods described earlier. From the bottom-left stacked bar chart of Figure 6, we observed that anosmia/ageusia were relatively low in order in

the frequency ranking (ie, 11th). As in Sarker et al's [4] ranking, cough came second in our data set's frequency ranking.

Finally, the bottom-right chart in Figure 6 refers to the comparison between our data set's frequency and importance rankings for the corresponding symptoms. We observed that anxiety ranked 4th in the frequency list, while it was low (ie, 23rd) in the importance ranking.

Figure 6. Feature comparison between our most important features and Sarker et al's [4] most frequent symptoms (top row) and between our most important features and our most frequent symptoms (bottom row). The feature importance rankings are obtained from an SVM linear kernel using the symptom-only vector representation. SVM: support vector machine.



Limitations

It is worth reiterating that social media posts, which are known to be noisy, are not on a par with the consultation that a patient would have with a doctor. We stress that the aim of this study is to extract useful information at a population level, rather than to provide an actionable decision for an individual via social media posts. Our manually annotated data set has 2 main limitations. First, having only 3 experts limited the quality of our labeling, although we deem this study to be a proof of concept. A larger number of experts, including more senior doctors, would be beneficial in a follow-up study. The robustness of our results could be further improved by both increasing the size of our data set and introducing posts from several alternate sources. Given that the posts come from social media, it is not clear whether the results could be used as such in a diagnostic system, without combining them with actual consultations. However, it is worth noting that medical social media, such as the posts we used herein, may uncover novel information regarding COVID-19.

Conclusion

The coronavirus pandemic has drawn a spotlight on the need to develop automated processes to provide additional information to researchers, health professionals, and decision makers. Medical social media comprises a rich resource of timely information that could fit this purpose. We have demonstrated that it is possible to take an approach that aims at the detection of COVID-19 using an automated triage and diagnosis system in order to augment public health surveillance systems, despite the heterogeneous nature of typical social media posts. The outputs from such an approach could be used to indicate the severity and estimate the prevalence of the disease in the population.

Authors' Contributions

All authors were involved in the design of the work. The first author wrote the code. The first 3 authors drafted the paper, and all authors critically revised the article.

Conflicts of Interest

None declared.

References

1. Overview: Coronavirus (COVID-19). URL: https://web.archive.org/web/20200316223405/https://www.nhs.uk/conditions/coronavirus-covid-19/ [accessed 2021-07-07]
2. Obeid JS, Davis M, Turner M, Meystre SM, Heider PM, O'Bryan EC, et al. An artificial intelligence approach to COVID-19 infection risk assessment in virtual visits: a case report. J Am Med Inform Assoc 2020 Aug 01;27(8):1321-1325 [FREE Full text] [doi: 10.1093/jamia/ocaa105] [Medline: 32449766]
3. Schwab P, DuMont Schütte A, Dietz B, Bauer S. Clinical predictive models for COVID-19: systematic study. J Med Internet Res 2020 Oct 06;22(10):e21439 [FREE Full text] [doi: 10.2196/21439] [Medline: 32976111]
4. Sarker A, Lakamana S, Hogg-Bremer W, Xie A, Al-Garadi MA, Yang Y. Self-reported COVID-19 symptoms on Twitter: an analysis and a research resource. J Am Med Inform Assoc 2020 Aug 01;27(8):1310-1315 [FREE Full text] [doi: 10.1093/jamia/ocaa116] [Medline: 32620975]
5. Qin L, Sun Q, Wang Y, Wu K, Chen M, Shia B, et al. Prediction of number of cases of 2019 novel coronavirus (COVID-19) using social media search index. Int J Environ Res Public Health 2020 Mar 31;17(7):2365 [FREE Full text] [doi: 10.3390/ijerph17072365] [Medline: 32244425]
6. Aramaki E, Maskawa S, Morita M. Twitter catches the flu: detecting influenza epidemics using Twitter. Proc EMNLP 2011:1568-1576.
7. Hu H, Wang H, Wang F, Langley D, Avram A, Liu M. Prediction of influenza-like illness based on the improved artificial tree algorithm and artificial neural network. Sci Rep 2018 Mar 20;8(1):4895 [FREE Full text] [doi: 10.1038/s41598-018-23075-1] [Medline: 29559649]
8. Norman G. Likert scales, levels of measurement and the "laws" of statistics. Adv Health Sci Educ Theory Pract 2010 Dec 10;15(5):625-632. [doi: 10.1007/s10459-010-9222-y] [Medline: 20146096]
9. Drucker H, Burges C, Kaufman L, Smola A, Vapnik V. Support vector regression machines. 1996 Presented at: NIPS'96: Proceedings of the 9th International Conference on Neural Information Processing Systems. Dec Presented atth International Conference on Neural Information Processing Systems; December 1996; Denver, CO p. 2-5. [doi: 10.1109/iconip.2002.1198219]
10. Marsland S. Machine Learning: An Algorithmic Perspective. 2nd ed. Boca Raton, FL: Chapman & Hall/CRC; 2014.
11. Arrieta A, García-Prado A, González P, Pinto-Prades JL. Risk attitudes in medical decisions for others: an experimental approach. Health Econ 2017 Dec 29;26 Suppl 3:97-113. [doi: 10.1002/hec.3628] [Medline: 29285873]
12. Sutton C. An introduction to conditional random fields. Found Trends Mach Learn 2012;4(4):267-373. [doi: 10.1561/2200000013]
13. Bach N, Badaskar S. A review of relation extraction. Lit Rev Lang Stat II 2007;2:15.
14. Edo-Osagie O, De La Iglesia B, Lake I, Edeghere O. A scoping review of the use of Twitter for public health research. Comput Biol Med 2020 Jul;122:103770 [FREE Full text] [doi: 10.1016/j.compbiomed.2020.103770] [Medline: 32502758]
15. Zens M, Brammertz A, Herpich J, Südkamp N, Hinterseer M. App-based tracking of self-reported COVID-19 symptoms: analysis of questionnaire data. J Med Internet Res 2020 Sep 09;22(9):e21956 [FREE Full text] [doi: 10.2196/21956] [Medline: 32791493]
16. Menni C, Valdes AM, Freidin MB, Sudre CH, Nguyen LH, Drew DA, et al. Real-time tracking of self-reported symptoms to predict potential COVID-19. Nat Med 2020 Jul 11;26(7):1037-1040 [FREE Full text] [doi: 10.1038/s41591-020-0916-2] [Medline: 32393804]
17. Mizrahi B, Shilo S, Rossman H, Kalkstein N, Marcus K, Barer Y, et al. Longitudinal symptom dynamics of COVID-19 infection. Nat Commun 2020 Dec 04;11(1):6208 [FREE Full text] [doi: 10.1038/s41467-020-20053-y] [Medline: 33277494]
18. Zimmerman RK, Balasubramani GK, Nowalk MP, Eng H, Urbanski L, Jackson ML, et al. Classification and Regression Tree (CART) analysis to predict influenza in primary care patients. BMC Infect Dis 2016 Sep 22;16(1):503 [FREE Full text] [doi: 10.1186/s12879-016-1839-x] [Medline: 27659721]
19. Brinati D, Campagner A, Ferrari D, Locatelli M, Banfi G, Cabitza F. Detection of COVID-19 infection from routine blood exams with machine learning: a feasibility study. J Med Syst 2020 Jul 01;44(8):135 [FREE Full text] [doi: 10.1007/s10916-020-01597-4] [Medline: 32607737]

20. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. BMJ 2020 Apr 07;369:m1328 [FREE Full text] [doi: 10.1136/bmj.m1328] [Medline: 32265220]
21. Judson T, Odisho A, Neinstein A, Chao J, Williams A, Miller C, et al. Rapid design and implementation of an integrated patient self-triage and self-scheduling tool for COVID-19. J Am Med Inform Assoc 2020 Jun 01;27(6):860-866 [FREE Full text] [doi: 10.1093/jamia/ocaa051] [Medline: 32267928]
22. Liu Y, Wang Z, Ren J, Tian Y, Zhou M, Zhou T, et al. A COVID-19 risk assessment decision support system for general practitioners: design and development study. J Med Internet Res 2020 Jun 29;22(6):e19786 [FREE Full text] [doi: 10.2196/19786] [Medline: 32540845]
23. Diagnosis of COVID-19 and its clinical spectrum AI and data science supporting clinical decision (from 28th Mar to 3rd Apr). Einstein Data4u. URL: https://www.kaggle.com/einsteindata4u/covid19 [accessed 2021-02-24]
24. Wang K, Zuo P, Liu Y, Zhang M, Zhao X, Xie S, et al. Clinical and laboratory predictors of in-hospital mortality in patients with coronavirus disease-2019: a cohort study in Wuhan, China. Clin Infect Dis 2020 Nov 19;71(16):2079-2088 [FREE Full text] [doi: 10.1093/cid/ciaa538] [Medline: 32361723]
25. Vaid A, Somani S, Russak AJ, De Freitas JK, Chaudhry FF, Paranjpe I, et al. Machine learning to predict mortality and critical events in a cohort of patients with COVID-19 in New York City: model development and validation. J Med Internet Res 2020 Nov 06;22(11):e24018 [FREE Full text] [doi: 10.2196/24018] [Medline: 33027032]
26. Wagner T, Shweta F, Murugadoss K, Awasthi S, Venkatakrishnan A, Bade S. Augmented curation of clinical notes from a massive EHR system reveals symptoms of impending COVID-19 diagnosis. Elife 2020:9. [doi: 10.7554/elife.58227]
27. Wang J, Abu-El-Rub N, Gray J, Pham H, Zhou Y, Manion F, et al. COVID-19 SignSym: a fast adaptation of a general clinical NLP tool to identify and normalize COVID-19 signs and symptoms to OMOP common data model. J Am Med Inform Assoc 2021 Jun 12;28(6):1275-1283 [FREE Full text] [doi: 10.1093/jamia/ocab015] [Medline: 33674830]
28. López-Úbeda P, Díaz-Galiano MC, Martín-Noguerol T, Luna A, Ureña-López LA, Martín-Valdivia MT. COVID-19 detection in radiological text reports integrating entity recognition. Comput Biol Med 2020 Dec;127:104066 [FREE Full text] [doi: 10.1016/j.compbiomed.2020.104066] [Medline: 33130435]
29. Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. Nucleic Acids Res 2004 Jan 01;32(Database issue):D267-D270 [FREE Full text] [doi: 10.1093/nar/gkh061] [Medline: 14681409]
30. Mikolov T, Sutskever I, Chen K, Corrado G, Dean J. Distributed representations of words and phrases and their compositionality. Proc NIPS 2013;2:3111-3119. [doi: 10.5040/9781474284974.00399]
31. Egton Medical Information Systems Limited. Coronavirus (COVID-19). Patient. URL: https://patient.info/forums/discuss/browse/coronavirus-covid-19--4541 [accessed 2022-02-18]
32. de Vet HCW, Mokkink LB, Terwee CB, Hoekstra OS, Knol DL. Clinicians are right not to like Cohen's κ. BMJ 2013 Apr 12;346(apr12 1):f2125-f2125. [doi: 10.1136/bmj.f2125] [Medline: 23585065]
33. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. J Clin Epidemiol 1990 Jan;43(6):543-549. [doi: 10.1016/0895-4356(90)90158-l]
34. Gwet K. Computing inter-rater reliability and its variance in the presence of high agreement. Br J Math Stat Psychol 2008;61(Pt 1):29-48. [doi: 10.1348/000711006x126600]
35. Wongpakaran N, Wongpakaran T, Wedding D, Gwet KL. A comparison of Cohen's kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. BMC Med Res Methodol 2013 Apr 29;13(1):61 [FREE Full text] [doi: 10.1186/1471-2288-13-61] [Medline: 23627889]
36. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]
37. Cunningham H, Maynard DK. Text Processing with GATE (Version 6). CA: Gateway Press; 2011.
38. Hasan A, Levene M, Weston D. Learning structured medical information from social media. J Biomed Inform 2020 Oct;110:103568 [FREE Full text] [doi: 10.1016/j.jbi.2020.103568] [Medline: 32942027]
39. Manning CH. Foundations of Statistical Natural Language Processing. 4th ed. Cambridge, MA: MIT Press; 2001.
40. python-crfsuite. URL: https://python-crfsuite.readthedocs.io/en/latest/ [accessed 2022-02-18]
41. Okazaki N. CRFsuite: A Fast Implementation of Conditional Random Fields (CRFs). URL: http://www.chokkan.org/software/crfsuite/ [accessed 2022-02-18]
42. Chang C, Lin C. LIBSVM. ACM Trans Intell Syst Technol 2011 Apr 01;2(3):1-27. [doi: 10.1145/1961189.1961199]
43. Abraham A, Pedregosa F, Eickenberg M, Gervais P, Mueller A, Kossaifi J, et al. Machine learning for neuroimaging with scikit-learn. Front Neuroinform 2014;8:14 [FREE Full text] [doi: 10.3389/fninf.2014.00014] [Medline: 24600388]
44. Support Vector Machines. URL: https://scikit-learn.org/stable/modules/svm.html [accessed 2022-02-18]
45. Weston J, Mukherjee S, Chapelle O, Pontil M, Poggio T, Vapnik V. Feature selection for SVMs. Adv Neural Inf Process Syst 2000;13:668-674.

Abbreviations

BPOC: body part, organ, or organ component

CRF: conditional random field
CT: computed tomography
EHR: electronic health record
GP: general physician
GTP: ground-truth probability
LR: logistic regression
NLP: natural language processing
RB: rule based
RBF: radial basis function
SVM: support vector machine
SVR: support vector regression
SYM: symptoms
XGB: gradient boosting

Edited by A Mavragani; submitted 13.05.21; peer-reviewed by A Trifan, S Doan; comments to author 03.07.21; revised version received 09.07.21; accepted 05.02.22; published 28.02.22.

Please cite as:
Hasan A, Levene M, Weston D, Fromson R, Koslover N, Levene T
Monitoring COVID-19 on Social Media: Development of an End-to-End Natural Language Processing Pipeline Using a Novel Triage and Diagnosis Approach
J Med Internet Res 2022;24(2):e30397
URL: https://www.jmir.org/2022/2/e30397
doi:10.2196/30397
PMID:

Original Paper

COVID-19 Vaccine Tweets After Vaccine Rollout: Sentiment–Based Topic Modeling

Luwen Huangfu[1,2], PhD; Yiwen Mo[1], MS; Peijie Zhang[3], MS; Daniel Dajun Zeng[3,4], PhD; Saike He[3,4], PhD

[1]Fowler College of Business, San Diego State University, San Diego, CA, United States

[2]Center for Human Dynamics in the Mobile Age, San Diego State University, San Diego, CA, United States

[3]The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China

[4]University of Chinese Academy of Sciences, Beijing, China

Corresponding Author:
Saike He, PhD
The State Key Laboratory of Management and Control for Complex Systems
Institute of Automation
Chinese Academy of Sciences
95 Zhongguancun East Road, Haidian District
Beijing, 100190
China
Phone: 86 (010)82544537
Email: saike.he@ia.ac.cn

Abstract

Background: COVID-19 vaccines are one of the most effective preventive strategies for containing the pandemic. Having a better understanding of the public's conceptions of COVID-19 vaccines may aid in the effort to promptly and thoroughly vaccinate the community. However, because no empirical research has yet fully explored the public's vaccine awareness through sentiment–based topic modeling, little is known about the evolution of public attitude since the rollout of COVID-19 vaccines.

Objective: In this study, we specifically focused on tweets about COVID-19 vaccines (Pfizer, Moderna, AstraZeneca, and Johnson & Johnson) after vaccines became publicly available. We aimed to explore the overall sentiments and topics of tweets about COVID-19 vaccines, as well as how such sentiments and main concerns evolved.

Methods: We collected 1,122,139 tweets related to COVID-19 vaccines from December 14, 2020, to April 30, 2021, using Twitter's application programming interface. We removed retweets and duplicate tweets to avoid data redundancy, which resulted in 857,128 tweets. We then applied sentiment–based topic modeling by using the compound score to determine sentiment polarity and the coherence score to determine the optimal topic number for different sentiment polarity categories. Finally, we calculated the topic distribution to illustrate the topic evolution of main concerns.

Results: Overall, 398,661 (46.51%) were positive, 204,084 (23.81%) were negative, 245,976 (28.70%) were neutral, 6899 (0.80%) were highly positive, and 1508 (0.18%) were highly negative sentiments. The main topics of positive and highly positive tweets were planning for getting vaccination (251,979/405,560, 62.13%), getting vaccination (76,029/405,560, 18.75%), and vaccine information and knowledge (21,127/405,560, 5.21%). The main concerns in negative and highly negative tweets were vaccine hesitancy (115,206/205,592, 56.04%), extreme side effects of the vaccines (19,690/205,592, 9.58%), and vaccine supply and rollout (17,154/205,592, 8.34%). During the study period, negative sentiment trends were stable, while positive sentiments could be easily influenced. Topic heatmap visualization demonstrated how main concerns changed during the current widespread vaccination campaign.

Conclusions: To the best of our knowledge, this is the first study to evaluate public COVID-19 vaccine awareness and awareness trends on social media with automated sentiment–based topic modeling after vaccine rollout. Our results can help policymakers and research communities track public attitudes toward COVID-19 vaccines and help them make decisions to promote the vaccination campaign.

(J Med Internet Res 2022;24(2):e31726) doi:10.2196/31726

KEYWORDS

COVID-19; COVID-19 vaccine; sentiment evolution; topic modeling; social media; text mining

Introduction

Background

COVID-19 vaccines are one of the most effective preventive strategies for containing the pandemic and restoring normal life [1]. The outcomes of this strategy highly depend on vaccination coverage, wherein herd immunity requires at least 70% of the population to be immune, depending on how contagious the COVID-19 variant in question is and how effective the vaccine is [2]. However, such a high rate of vaccination cannot be reached without the cooperation of the general public [3-5]. In general, there are a variety of factors that may negatively impact how the public perceives and reacts to these vaccines. Such barriers may stem from conspiracy theories [6], general hesitancy toward vaccines [4], and doubts regarding new mRNA vaccine technologies [7]. Infodemic management, that is, managing information overload, including false or misleading information [8], should be used during the COVID-19 pandemic, by listening to community concerns, preventing the spread of misleading information [9], and examining the human factors contributing to COVID-19 transmission [10]. Thus, to promote vaccine awareness and facilitate vaccine rollout, it is imperative to gain a timely understanding of the public's attitude toward vaccination and develop tailored communication strategies to address their concerns.

Generally, characterizing public vaccine attitudes as part of public health surveillance can be achieved via social media–based text mining or other traditional methodologies, such as conducting surveys or experiments. Social media–based text mining has become increasingly popular because of its effectiveness and efficiency; the major merit of this big data analysis is that it addresses several of the limitations of traditional methodologies, such as the inability to track real-time trends [4,11]. Public health monitoring on social media has proven to be a powerful tool for analyzing public health discussions on a variety of topics, such as pandemics and vaccination [12-24]. Such work has been conducted for the COVID-19 pandemic (Multimedia Appendix 1). However, because of the rapid COVID-19 vaccine rollout, dedicated social media–based sentiment analysis studies on vaccine awareness have just started to emerge [3,22-24]. Some of these studies [3,22] relied on natural language processing techniques to conduct large-scale sentiment analysis about vaccines, while others [23,24] investigated vaccination hesitancy using manual content analysis, but overall, these studies lacked either the capability to automatically track public attitudes (in manual content analysis) or a comprehensive view of both topics and associated sentiments. Furthermore, exploring the public sentiment and concern evolution throughout the current vaccination campaign may allow policymakers to make timely and informed decisions to encourage vaccination.

Study Objectives

We aimed to combine sentiment analysis and topic modeling in order to address the following research questions: What are the general sentiments on COVID-19 vaccines? What are the topics that shape the sentiments? How do concerns (ie, topics with negative sentiments) evolve over time?

Methods

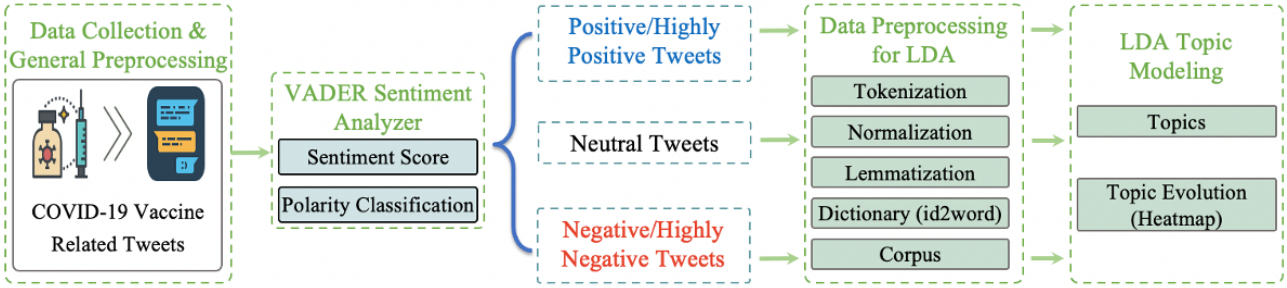
Data Collection

We collected COVID-19 vaccine–related tweets containing a variety of predefined hashtags, including #CovidVaccine, #GetVaccinated, #covid19vaccine, #vaccination, #AstraZeneca, #Johnson & Johnson, #Pfizer and #Moderna, from December 14, 2020 (after the first COVID-19 vaccine in the world was approved) to April 30, 2021. We collected 1,122,139 tweets (Table 1). To avoid data redundancy, we removed retweets and duplicate tweets, and we focused on tweets in English (Figure 1). After data preprocessing, the data set contained 857,128 tweets.

Table 1. Tweet hashtags.

Hashtag	Tweets (N=1,122,139), n
#CovidVaccine	345,537
#GetVaccinated	73,817
#covid19vaccine	130,043
#vaccination	132,327
#AstraZeneca	126,954
#Johnson & Johnson	211,731
#Pfizer	61,979
#Moderna	39,751

Figure 1. Data processing workflow. LDA: latent Dirichlet allocation; VADER: Valence Aware Dictionary for Sentiment Reasoning.



Sentiment Analysis

We used the Valence Aware Dictionary for Sentiment Reasoning (VADER) lexicon for analysis. During preprocessing, we did not remove the hashtag content because it often contained meaningful information such as the brand of the vaccine. VADER is a rule–based sentiment analysis tool that has been proven to perform as well as or even better than other sentiment analysis tools on social media texts in most cases, since it is specifically attuned to sentiments expressed on social media [25]. Generally, VADER produces 4 scores: positive, neutral, negative, and compound scores. Positive, neutral, and negative scores each represent the proportion of words that fall into the given category. The compound score is calculated by summing the valence scores of each word in the lexicon, adjusting the value according to heuristic rules, and normalizing between −1 and +1 [25]. The compound score is a useful metric for measuring the sentiment of each given text in a single dimension.

We classified each tweet into 1 of 5 groups (Table 2), based on compound, positive, and negative score distributions—highly positive (compound score >0.001 and positive sentiment score >0.5), positive (compound score >0.001 and positive sentiment score <0.5), highly negative (compound score <0.001 and negative sentiment score >0.5), and negative (compound score <0.001 and negative sentiment score <0.5), and neutral (if none of the conditions was satisfied).

Table 2. Sentiment polarity examples.

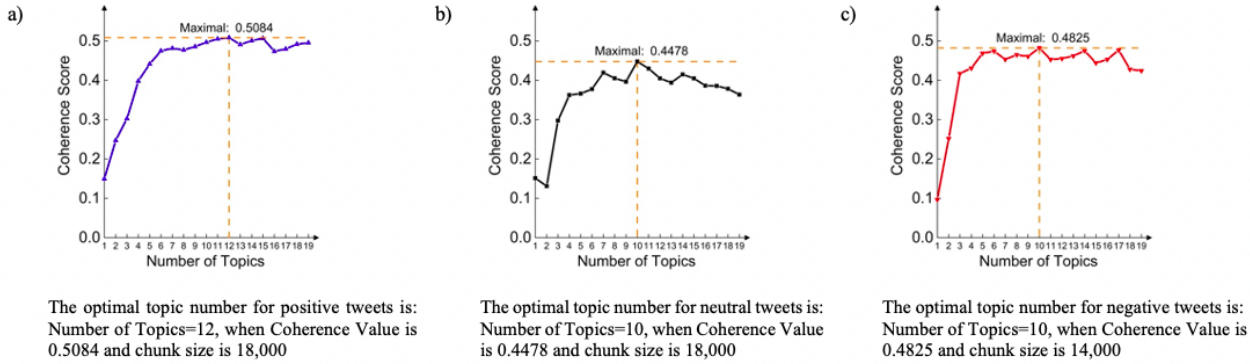
Sentiment polarity	Example
Highly positive	“thank god vaccination vaccinessavelives vaccineswork”
Positive	“it s an exciting day with the arrival of the first coronavirusvaccine it gives me great hope for 2021 covid19vaccine”
Highly negative	“it s fake you re all stupid covidvaccine”
Negative	“how do we know that after 6 9 months there are no adverse effects of the vaccine or that it s ineffective and what s the response if in the event these emergency approvals have larger ramifications any mechanism being put together covid_19 covid19vaccine”
Neutral	“help is on the way 1st doses of covid19vaccine arrived in north carolina initial vaccine supply is limited and will go to a small number of public health and hospital workers at high risk of exposure more doses are on the way but until then practice your 3ws”

Topic Modeling

Latent Dirichlet allocation (LDA), as a popular and well-established approach for topic analysis [26], is a three-level hierarchical Bayesian model that relies on the bag-of-words model [27]. LDA generates a probability distribution for the text corpus; it assumes that each topic can be characterized by a distribution of words. The number of topics is a key parameter of the LDA model. To prevent the misclassification of other topics into vaccine and nonvaccine topics, we removed some vaccine-related keywords, including “vaccine,” “vaccines,” “vaccination,” “covidvaccine,” and “covid.” This data preprocessing decision is also well supported by experimental results, which suggested that up to 96% of tweets were classified into one main topic with less meaningful information without removal of specific words.

To determine the optimal number of topics with favorable model performance, we used a coherence score; however, because the number of samples for highly positive and negative groups were small, we combined positive and highly positive groups (into a positive group) and negative and highly negative groups (into a negative group). Then, we applied topic modeling algorithms on 3 groups: positive, neutral, and negative. We used the topic coherence value to measure the modeling performance. Since the data set was very large, the experiments were run under the server environment with C5 computing type series IV 64-core CPU and 128 GB RAM. Then, based on the performance, we selected the optimal number of topics for each polarity group. The optimal topic numbers for positive, neutral, and negative were 12, 10, and 10, respectively (Figure 2).

Figure 2. Model performance for topic numbers for (a) positive, (b) neutral, and (c) negative tweets.



Results

Sentiment Analysis

Overall, positive sentiment was stronger than negative sentiment (Figure 3 and Figure 4). Notably, there was a sharp decline in the positive score around April 13, 2020 (Figure 3), which appeared to coincide with news released on that date: The US Federal Drug Administration (FDA) and Centers for Disease Control (CDC) called for a pause on the use of the Johnson & Johnson vaccine after discovering "extremely rare" cases of blood clots [28], and the number of tweets about the Johnson & Johnson vaccine peaked, reaching 23,729 tweets, which affect the average sentiment.

There were 6899 highly positive tweets, 398,661 positive tweets, 245,976 neutral tweets, 204,084 negative tweets, and 1508 highly negative tweets (Figure 5).

Figure 3. Overall daily average sentiment score.

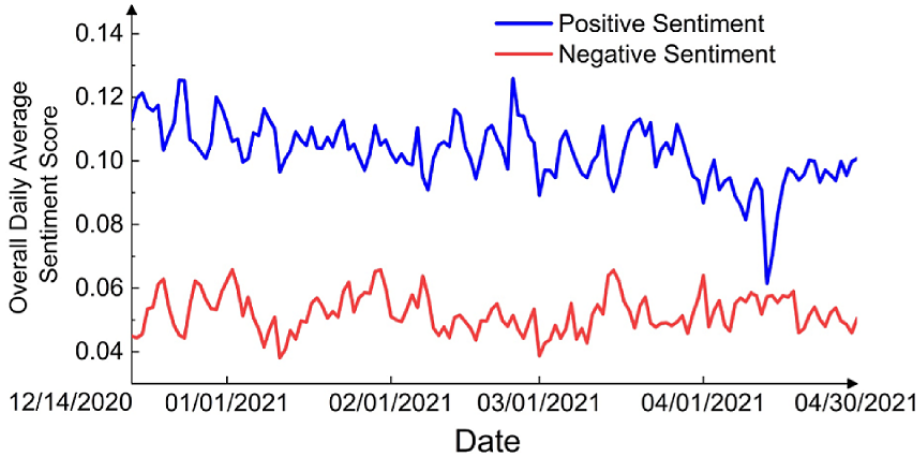


Figure 4. Overall sentiment trend.

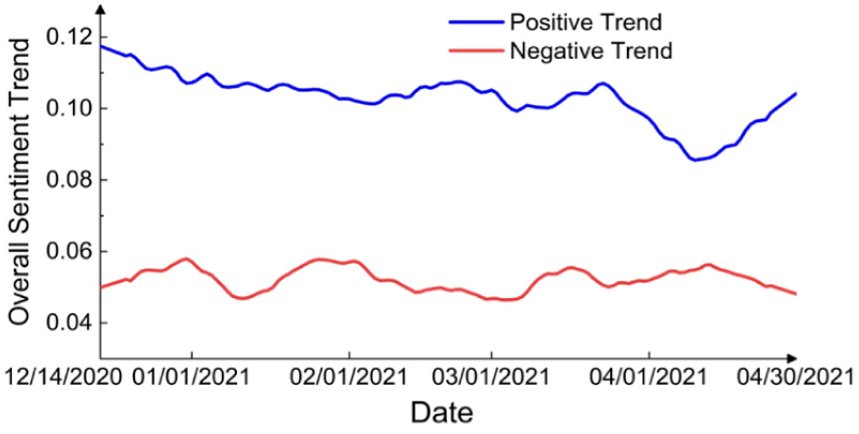
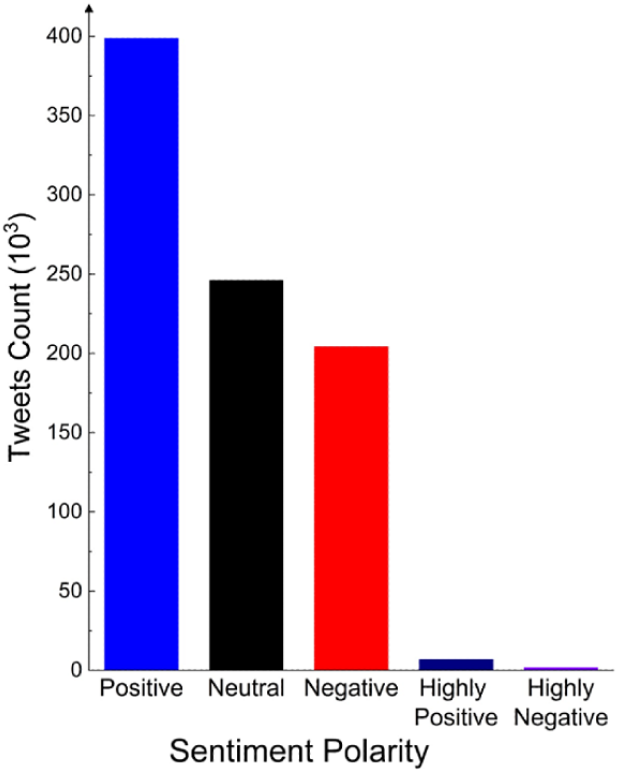


Figure 5. Sentiment polarity category distribution.



The percentage of negative sentiments was stable (Figure 6), but the percentage of positive sentiments decreased by month, and the percentage of neutral sentiments increased by month. Positive sentiment likely decreased due to the pause in the use of the Johnson & Johnson and AstraZeneca vaccinations in late March and April 2021 [28]. The neutral sentiment trend moved opposite to the positive sentiment trend.

Figure 6. Sentiment polarity distribution by month.

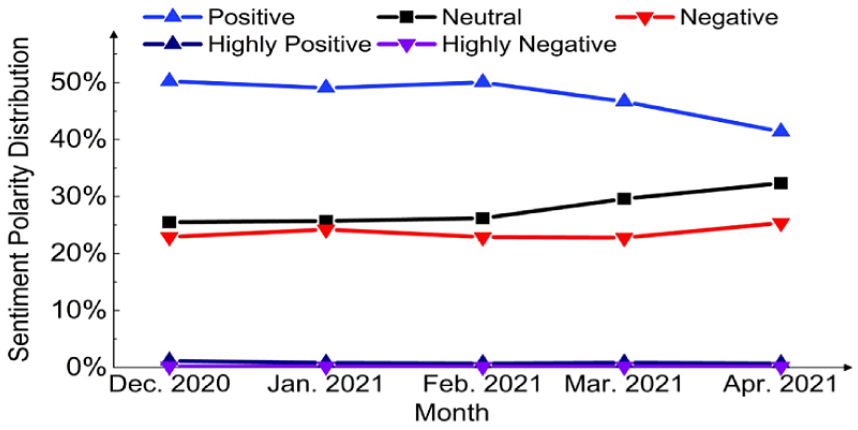


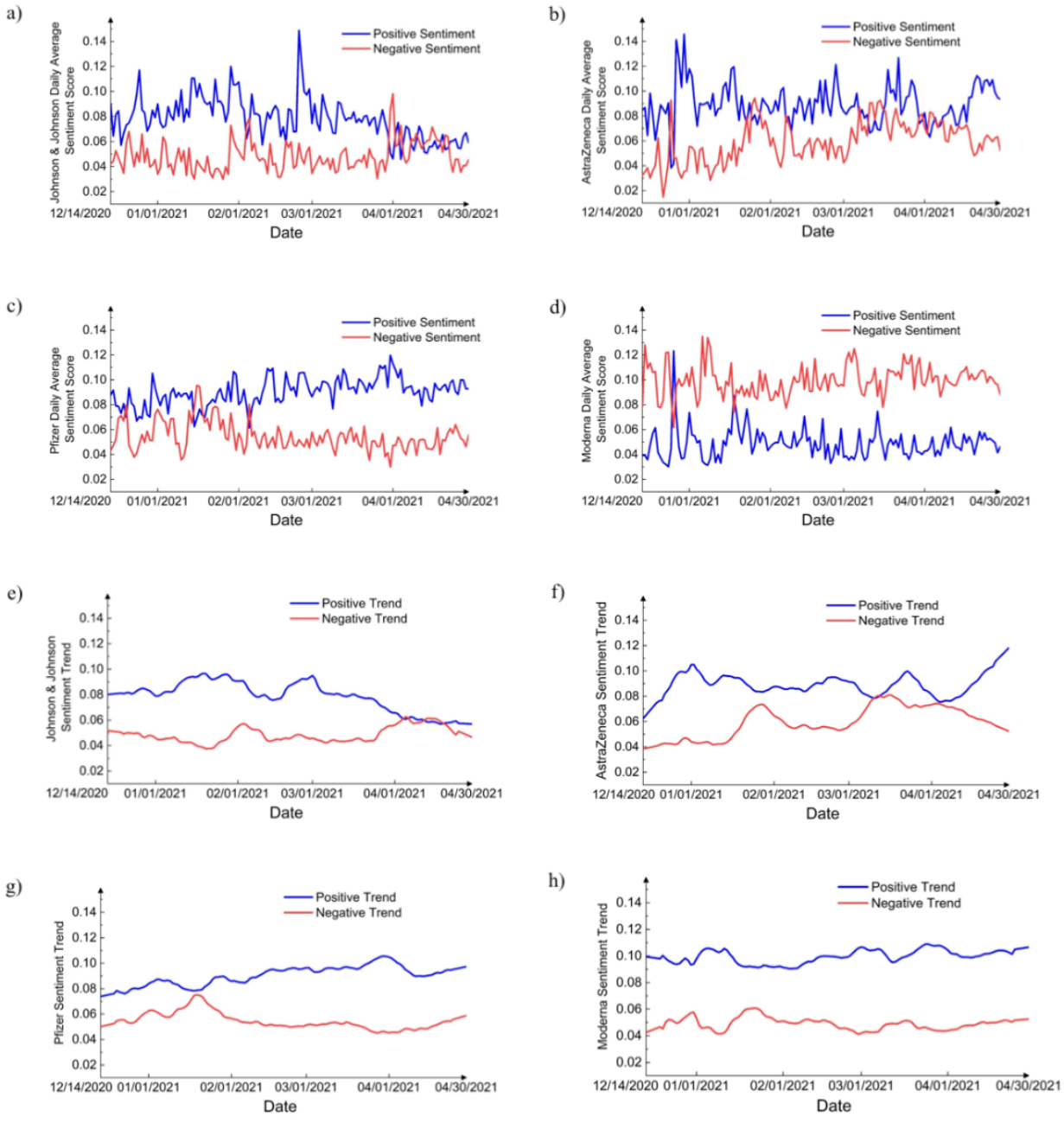
Figure 7 shows word clouds with profanities removed for highly positive, highly negative, positive, and negative tweets. Except for “vaccine” and “COVID,” which exhibited the highest frequency, the most common positive words in the highly positive group were “great,” “happy,” and “love”; the most common negative words in the highly negative group were “kill,” “bad,” and “death”; the most common positive words in the positive group were “thank,” “like,” and “health”; and the most common negative words in the negative group were “death,” “clot,” and “risk.”

Figure 7. Common words for (a) highly positive, (b) highly negative, (c) positive, and (d) negative tweets.

Additionally, the names of COVID-19 vaccine manufacturers *Johnson & Johnson* and *AstraZeneca* exhibited a high frequency in the negative groups.

Figure 8 shows that positive sentiment and negative sentiment scores changed daily for each vaccine and positive sentiment was stronger than negative sentiment; however, for Johnson & Johnson and AstraZeneca vaccines, the average positive and negative curves were found to intersect frequently, and the differences were small. From March 11 to March 16, 2021, distribution of the AstraZeneca vaccine was suspended in Europe [29]; however, on March 18, 2021, use of the AstraZeneca vaccine resumed in Europe after a review was conducted by the European Medicines Agency [30], which may be why positive and negative sentiment curves intersect in March 2021 and positive sentiment increased soon afterward. On April 13, 2021, FDA and CDC paused the use of the Johnson & Johnson vaccine due to several reports claimed that Johnson & Johnson might be linked to a very rare serious type of blood clotting in the vaccinated individuals. This explains why the negative sentiment trend increased and positive sentiment trend decreased in April 2021, even surpassing that of positive sentiments. On April 23, 2021, the FDA and CDC lifted the pause, but the positive trend was stable and remained low, which reflected the public's concerns about the Johnson & Johnson and AstraZeneca vaccines.

Figure 8. Daily average positive and negative sentiment scores for (a) Johnson & Johnson, (b) AstraZeneca, (c) Pfizer, and (d) Moderna vaccines and sentiment trends for (e) Johnson & Johnson, (f) AstraZeneca, (g) Pfizer, and (h) Moderna vaccines.



For Pfizer and Moderna vaccines, positive and negative sentiment curves were found to intersect only in December 2020 and January 2021, and the sentiment trends were stable, which reflected public concerns in the beginning, when the vaccines were first approved, followed by increasing levels of confidence in the vaccines as more and more people became vaccinated.

Figure 9 shows the standard deviation of sentiments for each vaccine. For the Pfizer and Moderna vaccines, the standard deviation lines are flat, which means that the sentiments for these vaccines were very stable and did not exhibit much fluctuation. However, for Johnson & Johnson and AstraZeneca vaccines, the standard deviation of sentiments changed drastically over time. For instance, the standard deviation of the Johnson & Johnson vaccine decreased, implying a higher degree of consensus regarding this specific vaccine. However, the opposite was true for the AstraZeneca vaccine, and the increased sentiment variation indicated the attitudes toward it were found to be more divided over time.

Figure 10 shows the percentages of tweets for each vaccine in each sentiment polarity; the percentages in each sentiment group are very close to each other.

Figure 9. Daily standard deviation of sentiments for (a) Johnson & Johnson, (b) AstraZeneca, (c) Pfizer, and (d) Moderna vaccines.

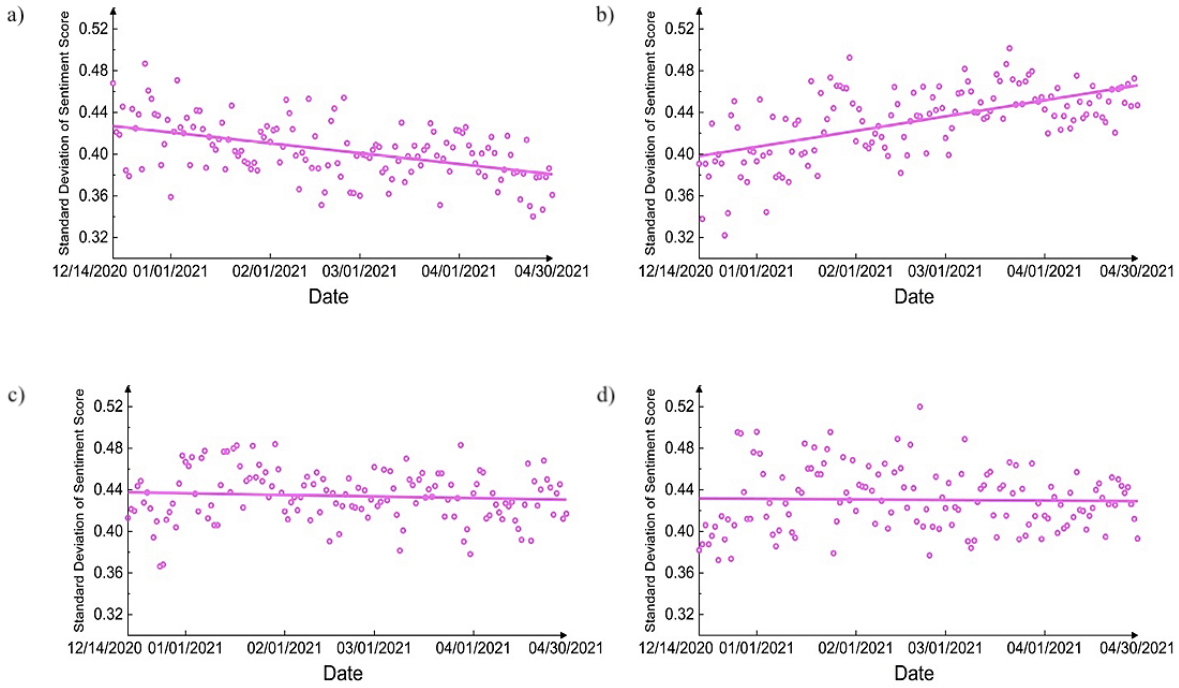
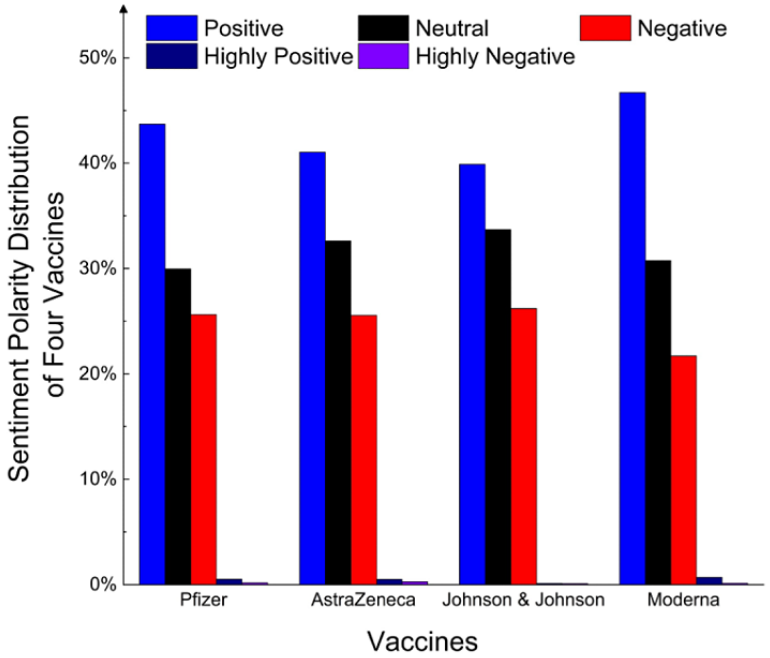


Figure 10. Sentiment polarity distributions for Pfizer, AstraZeneca, Johnson & Johnson, and Moderna vaccines.



Topic Modeling

Positive Topics

Topics suggested that people felt happy and grateful that a vaccine had been approved (Table 3), that it is important to get vaccinated, that they were thankful to the health care staff for their efforts, and that they were waiting to be eligible for vaccination.

Table 3. Top 5 positive (including highly positive) topics.

Topic ID	Tweets, n (%)	Keywords	Topic
POS_05	251,979 (62.13)	people, take, say, make, go, good, need, help, well, give	Planning for getting vaccination
POS_07	76,029 (18.75)	get, today, dose, first, feel, shoot, day, second, shot, be	Getting vaccinated
POS_09	21,127 (5.21)	share, read, important, health, join, question, public, information, community, concern	Vaccine information and knowledge
POS_11	14,286 (3.52)	thank, clinic, staff, support, team, volunteer, work, process, amazing, effort	Thanks for healthcare worker
POS_01	6,963 (1.72)	effective, risk, variant, pause, blood_clot, virus, benefit, less, rare, infection	Side effects

Neutral Topics

The main neutral topics were vaccination appointment (79,710/245,976, 32.41%) and getting vaccinated (40,532/245,976, 16.48%) (Table 4). Even though the topics were neutral, they revealed favorable attitudes toward COVID-19 vaccines. In addition, 12.77% (31,409/245,976) of neutral tweets demonstrated that people felt some hesitancy toward receiving the vaccine or that they need more time to think and make a decision.

Table 4. Top 5 neutral topics.

Topic ID	Tweets, n (%)	Keywords	Topic
NEU_05	79,710 (32.41)	get, today, appointment, shoot, available, be, call, wait, come, schedule	Vaccination appointment
NEU_02	40,532 (16.48)	dose, first, receive, second, shot, pfizer, day, week, administer, fully	Getting vaccinated
NEU_09	31,409 (12.77)	say, take, go, people, time, still, need, rare, would, think	Vaccine hesitancy
NEU_03	17,156 (6.97)	update, read, find, late, live, news, check, watch, question, link	Vaccine news
NEU_06	17,129 (6.96)	may, start, age, year, week, open, next, eligible, site, begin	Vaccine eligibility

Negative Topics

Negative topics (Table 5) demonstrated the public's main concerns regarding COVID-19 vaccines. In general, the public mainly cared about the side effects of vaccines, including common side effects, such as soreness after receiving a vaccine, and serious adverse reactions, such as death. However, given the strict storage requirement, the vaccines' supply chain and rollout were the second most important issue that concerned the public. Other negative topics involved the vaccination appointment, coronavirus variants, vaccination for women and patients with cancer (people who are at high risk), fake news, and misinformation.

Table 5. Negative (including highly negative) topics.

Topic ID	Tweets, n (%)	Keywords	Topics
NEG_05	115,206 (56.04)	get, people, take, go, say, make, know, stop, need, still	Vaccine hesitancy
NEG_00	19,690 (9.58)	risk, death, case, report, blood_clot, rare, severe, low, receive, blood	Extreme side effects
NEG_06	17,154 (8.34)	government, country, pay, company, rollout, state, plan, fail, stock, supply	Vaccine supply and rollout
NEG_04	14,125 (6.87)	get, shoot, feel, arm, day, hour, today, shot, sore, second	Common side effects
NEG_07	10,248 (4.98)	appointment, wait, available, age, site, open, today, hospital, group, offer	Vaccination appointment
NEG_03	8080 (3.93)	use, emergency, say, suspend, break, astrazeneca, official, country, shortage, pause	AstraZeneca suspension
NEG_02	7100 (3.45)	dose, week, first, second, receive, next, day, ruin, delay, administer	Vaccine administration
NEG_09	6151 (2.99)	read, question, health, public, story, information, hesitancy, register, community, explain	Vaccine information and community
NEG_01	4471 (2.17)	pandemic, virus, new, fight, variant, lockdown, avoid, coronavirus, spread, restriction	Spread avoidance
NEG_08	3367 (1.64)	cause, cancer, clot, woman, trust, product, doctor, body, choice, damage	Extreme side effects on vulnerable groups

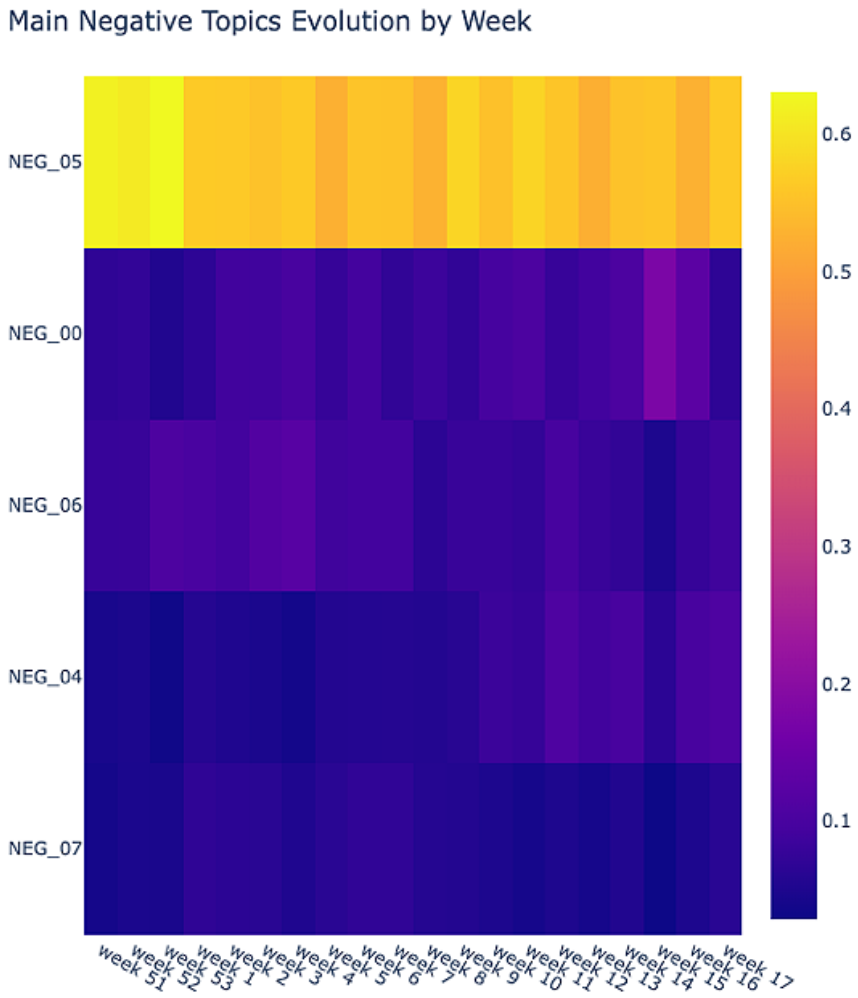
We found that 47.32% of the tweets (405,560/857,128), demonstrated positive (including highly positive) attitudes toward COVID-19 vaccines. The main topics included encouraging people to get vaccinated and conveying hope and gratitude for future life as a result of vaccine approval. Overall, 23.99% of the tweets (205,592/857,128) expressed negative (including highly negative) attitudes and concerns. The main

concerns regarding COVID-19 vaccines were side effects of vaccination, serious adverse reactions, and vaccine supply.

Topic Evolution

Side effects, such as pain at the injection site (ie, NEG_05) were discussed the most (of all negative topics) throughout the period (Figure 11). Moreover, with the increase in the number of people who received the vaccine, the discussion on side effects increased. Topics such as vaccine supply (ie, NEG_00) and extreme side effects (ie, NEG_06) were discussed less but a consistent amount throughout the period.

Figure 11. Heatmap of negative topic evolution. The x-axis represents the week in the year. Lighter colors correspond to topics that are discussed more.



Discussion

General Sentiments

Most sentiments toward COVID-19 vaccines were neutral and positive. Positive sentiment was stronger than negative sentiment throughout the period. Previous results from research conducted from March 1 to November 22, 2020 (before vaccines were available) [3] were similar—the dominant sentiments were positive and neutral; however, in this study, negative sentiment (205,592/857,128, 23.99%) was lower than that in [3] (30.57%). This suggests that after the COVID-19 vaccines became available, their effectiveness in reducing the risk of infection started to manifest in the real world, and people started having fewer doubts on social media toward vaccines. Vaccine trials, social media, and government interventions may contribute to alleviating public concerns [31].

Concerns and Topics That Shape Attitudes

By applying topic modeling to our data set, we found that the main topic in the positive and neutral domain was encouraging people to get vaccinated. In general, we discovered that vaccines are becoming widely accepted by the public as time passes. The main topic of our negative data set was the severe side effects of vaccination. When some social media outlets reported possible vaccination side effects, the concerns were discussed frequently on different social media platforms, such as Twitter, and possibly impacted individual decisions. Before vaccines were available, discussions on vaccines were centered around clinical trials and vaccine availability [12]. However, upon vaccine rollout, the concerns shifted dramatically to common side effects, which dominated the discussion throughout the study period (from December 14, 2020 to April 30, 2021). Hence, timely monitoring of the public attitude can help guide

public health officials to communicate more effectively with the public.

We also found that among the negative tweets, other than vaccine hesitancy, the main concerns regarding side effects (NEG_00 and NEG_04) were vaccine supply and rollout (NEG_06). This finding is consistent with those from previous studies [24,32,33]. For example, in a study on vaccination hesitancy in Canada [24], it was found that vaccination hesitancy stemmed from mistrust toward vaccine development, lack of knowledge about COVID-19 vaccines, and suspicion about political and authority figures who were not taking the vaccine. In another study [32] employing a questionnaire for the Israeli population, the results showed that the top 3 concerns regarding COVID-19 vaccines were quality control, side effects, and doubtful efficiency. Another survey conducted in the United States and Canada showed that vaccine rejection is very strongly related to vaccine benefits, vaccine safety, and unforeseen future effects [33]. Overall, our findings were similar—the top concerns were vaccine safety, side effects, vaccine supply, and government policy.

Changes by Month

Overall, it was observed that positive sentiment distribution decreased, neutral sentiment distribution increased, and negative sentiment distribution was stable. However, positive sentiment was dominant throughout the study period (December 14, 2020 to April 30, 2021). Positive sentiment decreased in March and April 2021, likely because of the extreme side effects (blood clotting) reported in the news for Johnson & Johnson and AstraZeneca vaccines. Use of the AstraZeneca vaccine was even stopped in Europe briefly [29], and the FDA and CDC called for a pause on the use of the Johnson & Johnson vaccine in the United States [28]. This may have caused positive sentiment to decrease, while neutral sentiment rather than negative sentiment increased, because people tended to feel neutral rather than very negative, toward such a pause.

In the very beginning, such side effects were extensively discussed. Some news outlets reported severe side effects, such as Bell palsy and even death [34], after receiving the vaccine, which seemed to coincide with more negative sentiments. Both Pfizer and Moderna vaccines are mRNA vaccines, which is a new type of vaccine that has not been used before [35]. This caused the general public to have concerns regarding the long-term side effects of these novel vaccines [7]. In the beginning, the lack of knowledge about COVID-19 and mRNA vaccines shaped the public's concerns. However, as more people were vaccinated over time, more people were able to observe how these vaccines helped steadily decrease the number of new cases and deaths per day as well as the hospitalization rates, implying that the pandemic is under control thanks to these vaccines. This in turn resulted in an increasing number of people seeking to become vaccinated, because extreme side effects are very rare and might be associated with misinformation and because the common side effects are regarded as tolerable.

Sentiment trend findings were consistent with those from a previous study [22] in which a vaccine acceptance experiment using Weibo Sina (a popular social media platform in China) demonstrated that positive attitudes were dominant, that the Chinese population were inclined to be positive about the side effects over time, and that one of the concerns that affects vaccine acceptance are misunderstandings about vaccination.

Limitations and Future Work

In this study, we mainly focused on textual information from the Twitter platform. However, users may be distributed among different social media platforms and different locations according to their usage, language, and preferences. Therefore, the methods used in our study can be extended to different social media platforms. It is also possible to use geographical filters on location information or to work on other languages to precisely differentiate between the significant issues and concerns among the different cultures or demographics.

Furthermore, our model can be extended to other research problems. For example, future studies should focus on negative tweets to determine whether misinformation exists or to identify misinformation on social media and propose suggestions for how to minimize the spread of such misinformation. Moreover, it may be plausible in the future to train a topic model with LDA and deep learning to forecast event topics and trends.

Conclusions

Our work profiles the spectrum of public sentiments toward vaccination and the main concerns underlying these views since the rollout of vaccines. These findings demonstrate the effectiveness of sentiment–based topic modeling in identifying topics and trends in polarity groups and in revealing the dynamic nature of public attitudes toward vaccination in the midst of evolving situations and changing public measures during the pandemic. Adding sentiment analysis and topic modeling when monitoring COVID-19 vaccine awareness can help researchers uncover time–based viewpoints underlying the dynamic public attitude toward vaccination on a large scale and devise tailored communication strategies to promote vaccination.

Acknowledgments

This study was supported by a San Diego State University Master Research Scholarship and by Research Funds from Fowler College of Business. We thank Professor David Banks from Duke University for providing helpful and constructive comments and suggestions.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Related work on sentiment analysis or topic modeling.
[DOCX File , 53 KB - jmir_v24i2e31726_app1.docx]

References

1. COVID-19 vaccines are effective. Centers for Disease Control and Prevention. 2021 Dec 23. URL: https://www.cdc.gov/coronavirus/2019-ncov/vaccines/effectiveness.html [accessed 2021-12-28]
2. D'Souza G, Dowdy D. Rethinking herd immunity and the covid-19 response endgame. Johns Hopkins Bloomberg School of Public Health. Sep 13. URL: https://publichealth.jhu.edu/2021/what-is-herd-immunity-and-how-can-we-achieve-it-with-covid-19 [accessed 2021-12-28]
3. Hussain A, Tahir A, Hussain Z, Sheikh Z, Gogate M, Dashtipour K, et al. Artificial intelligence-enabled analysis of public attitudes on Facebook and Twitter toward covid-19 vaccines in the United Kingdom and the United States: observational study. J Med Internet Res 2021 Apr 05;23(4):e26627 [FREE Full text] [doi: 10.2196/26627] [Medline: 33724919]
4. Lazarus JV, Ratzan SC, Palayew A, Gostin LO, Larson HJ, Rabin K, et al. A global survey of potential acceptance of a covid-19 vaccine. Nat Med 2021 Feb;27(2):225-228 [FREE Full text] [doi: 10.1038/s41591-020-1124-9] [Medline: 33082575]
5. Leischow SJ, Milstein B. Systems thinking and modeling for public health practice. Am J Public Health 2006 Mar;96(3):403-405. [doi: 10.2105/AJPH.2005.082842] [Medline: 16449572]
6. Romer D, Jamieson KH. Conspiracy theories as barriers to controlling the spread of COVID-19 in the U.S. Soc Sci Med 2020 Oct;263:113356 [FREE Full text] [doi: 10.1016/j.socscimed.2020.113356] [Medline: 32967786]
7. Hitti FL, Weissman D. Debunking mRNA vaccine misconceptions-an overview for medical professionals. Am J Med 2021 Jun;134(6):703-704 [FREE Full text] [doi: 10.1016/j.amjmed.2021.02.004] [Medline: 33737059]
8. Infodemic. World Health Organization. URL: https://www.who.int/health-topics/infodemic#tab=tab_1 [accessed 2021-12-28]
9. Eysenbach G. How to fight an infodemic: the four pillars of infodemic management. J Med Internet Res 2020 Jun 29;22(6):e21820 [FREE Full text] [doi: 10.2196/21820] [Medline: 32589589]
10. Serrano P, Huangfu L. CURVE4COVID: comprehensive understanding via representative variable exploration for COVID-19. In: Proceedings of 2021 Americas Conference on Information Systems. 2021 Presented at: Americas Conference on Information Systems; August 9-13; Online p. 1723-1733 URL: https://aisel.aisnet.org/amcis2021/healthcare_it/sig_health/27/
11. Li X, Zeng W, Li X, Chen H, Shi L, Li X, et al. CT imaging changes of corona virus disease 2019(COVID-19): a multi-center study in Southwest China. J Transl Med 2020 Apr 06;18(1):154 [FREE Full text] [doi: 10.1186/s12967-020-02324-w] [Medline: 32252784]
12. Wen A, Wang L, He H, Liu S, Fu S, Sohn S, et al. An aberration detection-based approach for sentinel syndromic surveillance of COVID-19 and other novel influenza-like illnesses. J Biomed Inform 2021 Jan;113:103660 [FREE Full text] [doi: 10.1016/j.jbi.2020.103660] [Medline: 33321199]
13. Li L, Zhang Q, Wang X, Zhang J, Wang T, Gao T, et al. Characterizing the propagation of situational information in social media during COVID-19 epidemic: a case study on Weibo. IEEE Trans Comput Soc Syst 2020 Apr;7(2):556-562. [doi: 10.1109/tcss.2020.2980007]
14. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. PLoS One 2010 Nov 29;5(11):e14118 [FREE Full text] [doi: 10.1371/journal.pone.0014118] [Medline: 21124761]
15. Signorini A, Segre AM, Polgreen PM. The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. PLoS One 2011 May 04;6(5):e19467 [FREE Full text] [doi: 10.1371/journal.pone.0019467] [Medline: 21573238]
16. Mutanga M, Abayomi A. Tweeting on COVID-19 pandemic in South Africa: LDA-based topic modelling approach. Africa J Sci Technol Innov Dev 2020 Oct 08:1-10. [doi: 10.1080/20421338.2020.1817262]
17. Oyebode O, Ndulue C, Adib A, Mulchandani D, Suruliraj B, Orji FA, et al. Health, Psychosocial, and Social Issues Emanating From the COVID-19 Pandemic Based on Social Media Comments: Text Mining and Thematic Analysis Approach. JMIR Med Inform 2021 Apr 06;9(4):e22734 [FREE Full text] [doi: 10.2196/22734] [Medline: 33684052]
18. Jang H, Rempel E, Roth D, Carenini G, Janjua NZ. Tracking COVID-19 Discourse on Twitter in North America: Infodemiology Study Using Topic Modeling and Aspect-Based Sentiment Analysis. J Med Internet Res 2021 Feb 10;23(2):e25431 [FREE Full text] [doi: 10.2196/25431] [Medline: 33497352]
19. Garcia K, Berton L. Topic detection and sentiment analysis in Twitter content related to COVID-19 from Brazil and the USA. Appl Soft Comput 2021 Mar 10;101(2):107057 [FREE Full text] [doi: 10.1016/j.asoc.2020.107057] [Medline: 33519326]
20. S.V. P, Ittamalla R. An analysis of attitude of general public toward COVID-19 crises – sentimental analysis and a topic modeling study. Inf Discov Deliv 2021 Feb 11;49(3):240-249 [FREE Full text] [doi: 10.1108/idd-08-2020-0097]
21. Abdulaziz M, Alotaibi A, Alsolamy M, Alabbas A. Topic based sentiment analysis for covid-19 tweets. Int J Adv Comput Sci Appl 2021;12(1). [doi: 10.14569/ijacsa.2021.0120172]

22. Yin F, Wu Z, Xia X, Ji M, Wang Y, Hu Z. Unfolding the determinants of covid-19 vaccine acceptance in China. J Med Internet Res 2021 Jan 15;23(1):e26089 [FREE Full text] [doi: 10.2196/26089] [Medline: 33400682]
23. Hou Z, Tong Y, Du F, Lu L, Zhao S, Yu K, et al. Assessing covid-19 vaccine hesitancy, confidence, and public engagement: a global social listening study. J Med Internet Res 2021 Jun 11;23(6):e27632 [FREE Full text] [doi: 10.2196/27632] [Medline: 34061757]
24. Griffith J, Marani H, Monkman H. COVID-19 vaccine hesitancy in Canada: content analysis of tweets using the theoretical domains framework. J Med Internet Res 2021 Apr 13;23(4):e26874 [FREE Full text] [doi: 10.2196/26874] [Medline: 33769946]
25. Hutto C, Gilbert E. VADER: a parsimonious rule-based model for sentiment analysis of social media text. In: Proceedings of the International AAAI Conference on Web and Social Media. 2014 Presented at: International AAAI Conference on Web and Social Media; June 1-4; Ann Arbor, Michigan p. 216-225 URL: https://ojs.aaai.org/index.php/ICWSM/article/view/14550
26. Jelodar H, Wang Y, Yuan C, Feng X, Jiang X, Li Y, et al. Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. arXiv 2021 Preprint posted online on Dec 6, 2018. [FREE Full text]
27. Blei DM, Ng AY, Jordan MI. Latent Dirichlet allocation. J Mach Learn Res 2003;3:993-1022 [FREE Full text]
28. US Food and Drug Administration. URL: https://www.fda.gov/news-events/press-announcements/joint-cdc-and-fda-statement-johnson-johnson-covid-19-vaccine [accessed 2021-12-30]
29. AstraZeneca shares slide as clotting reports lead Denmark to pause rollout of its vaccine. Fortune. URL: https://fortune.com/2021/03/11/astrazeneca-shares-slide-clotting-reports-denmark-pause-rollout/ [accessed 2021-12-30]
30. AstraZeneca's COVID vaccine gets all-clear from EU health agency following blood clot uproar. Fortune. URL: https://fortune.com/2021/03/18/astrazenecas-covid-vaccine-all-clear-eu-health-agency-blood-clot-uproar/ [accessed 2021-12-30]
31. Bavel JJV, Baicker K, Boggio PS, Capraro V, Cichocka A, Cikara M, et al. Using social and behavioural science to support COVID-19 pandemic response. Nat Hum Behav 2020 May;4(5):460-471. [doi: 10.1038/s41562-020-0884-z] [Medline: 32355299]
32. Dror AA, Eisenbach N, Taiber S, Morozov NG, Mizrachi M, Zigron A, et al. Vaccine hesitancy: the next challenge in the fight against COVID-19. Eur J Epidemiol 2020 Aug;35(8):775-779. [doi: 10.1007/s10654-020-00671-y] [Medline: 32785815]
33. Taylor S, Landry CA, Paluszek MM, Groenewoud R, Rachor GS, Asmundson GJG. A proactive approach for managing covid-19: the importance of understanding the motivational roots of vaccination hesitancy for SARS-CoV2. Front Psychol 2020;11:575950 [FREE Full text] [doi: 10.3389/fpsyg.2020.575950] [Medline: 33192883]
34. Buntz B. Is there a link between Bell's palsy and COVID-19 vaccines? Drug Discovery Trends. 2021 Mar 1. URL: https://www.drugdiscoverytrends.com/is-there-a-link-between-bells-palsy-and-covid-19-vaccines/ [accessed 2021-12-30]
35. Beyrer C. The long history of mRNA vaccines. Johns Hopkins Bloomberg School of Public Health. 2021 Oct 06. URL: https://publichealth.jhu.edu/2021/the-long-history-of-mrna-vaccines [accessed 2021-12-30]

Abbreviations

CDC: Centers for Disease Control
FDA: US Food and Drug Administration
LDA: latent Dirichlet allocation
VADER: Valence Aware Dictionary for Sentiment Reasoning

Edited by G Eysenbach, M Gisondi; submitted 02.07.21; peer-reviewed by C Liu, Z Hu, K Dashtipour; comments to author 23.07.21; revised version received 12.11.21; accepted 13.11.21; published 08.02.22.

Please cite as:
Huangfu L, Mo Y, Zhang P, Zeng DD, He S
COVID-19 Vaccine Tweets After Vaccine Rollout: Sentiment–Based Topic Modeling
J Med Internet Res 2022;24(2):e31726
URL: https://www.jmir.org/2022/2/e31726
doi:10.2196/31726
PMID:34783665

Original Paper

A Stanford Conference on Social Media, Ethics, and COVID-19 Misinformation (INFODEMIC): Qualitative Thematic Analysis

Michael A Gisondi[1*], MD; Daniel Chambers[2*]; Tatum Minh La[2*]; Alexa Ryan[2*]; Adyant Shankar[2*]; Athena Xue[2*]; Rachel Anne Barber[2*]

[1]Department of Emergency Medicine, Stanford School of Medicine, Palo Alto, CA, United States

[2]Stanford University, Palo Alto, CA, United States

[*]all authors contributed equally

Corresponding Author:
Michael A Gisondi, MD
Department of Emergency Medicine
Stanford School of Medicine
900 Welch Road
Suite 350
Palo Alto, CA, 94304
United States
Phone: 1 650 721 4023
Email: mgisondi@stanford.edu

Abstract

Background: The COVID-19 pandemic continues to challenge the world's population, with approximately 266 million cases and 5 million deaths to date. COVID-19 misinformation and disinformation led to vaccine hesitancy among the public, particularly in vulnerable communities, which persists today. Social media companies are attempting to curb the ongoing spread of an overwhelming amount of COVID-19 misinformation on their platforms. In response to this problem, the authors hosted *INFODEMIC: A Stanford Conference on Social Media and COVID-19 Misinformation* (INFODEMIC) to develop best practices for social media companies to mitigate online misinformation and disinformation.

Objective: The primary aim of this study was to develop recommendations for social media companies to address the COVID-19 infodemic. We report the methods used to execute the INFODEMIC conference, conference attendee engagement and analytics, and a qualitative thematic analysis of the conference presentations. The primary study outcomes were the identified themes and corresponding recommendations.

Methods: Using a constructivist paradigm, we conducted a thematic analysis of the 6-hour conference transcript to develop best practice recommendations. The INFODEMIC conference was the study intervention, the conference speakers were the study participants, and transcripts of their presentations were the data for this study. We followed the 6-step framework for thematic analysis described by Braun and Clarke. We also used descriptive statistics to report measures of conference engagement including registrations, viewership, post-conference asynchronous participation, and conference evaluations.

Results: A total of 26 participants spoke at the virtual conference and represented a wide array of occupations, expertise, and countries of origin. From their remarks, we identified 18 response categories and 4 themes: trust, equity, social media practices, and interorganizational partnerships. From these, a total of 16 best practice recommendations were formulated for social media companies, health care organizations, and the general public. These recommendations focused on rebuilding trust in science and medicine among certain communities, redesigning social media platforms and algorithms to reduce the spread of misinformation, improving partnerships between key stakeholders, and educating the public to critically analyze online information. Of the 1090 conference registrants, 587 (53.9%) attended the live conference, and another 9996 individuals viewed or listened to the conference recordings asynchronously. Conference evaluations averaged 8.9 (best=10).

Conclusions: Social media companies play a significant role in the COVID-19 infodemic and should adopt evidence-based measures to mitigate misinformation on their platforms.

(J Med Internet Res 2022;24(2):e35707) doi:10.2196/35707

KEYWORDS

COVID-19; infodemic; misinformation; disinformation; vaccine; social media; thematic analysis; qualitative

Introduction

The COVID-19 pandemic has taken a huge toll on the world. COVID-19 is currently the third leading cause of death after heart disease and cancer, with approximately 266 million cases and 5 million deaths to date [1,2]. More than 780,000 Americans have died from COVID-19, and new variants of the virus continue to emerge. However, after an initial year of staggering case numbers, we finally have observed a decrease in the number of new infections requiring hospitalization as more people get vaccinated against the virus [1]. As of December 2021, 60% of the US adult population had received 2 doses of the COVID-19 vaccine [3,4]. Still, this fell short of the US Centers for Disease Control and Prevention (CDC) goal of 70% vaccination of the US population by July 4, 2021 [5]. This disparity persists despite widespread vaccine access [3].

Throughout the COVID-19 pandemic, there has been a proliferation of both misinformation and disinformation about the virus, its origin, the vaccines, and potential treatments. Misinformation refers to inaccurate information disseminated without malice, while disinformation is the purposeful spread of inaccurate information with malicious intent (Table 1). Taken together, the escalation of inaccurate information surrounding the pandemic can be accurately described as the COVID-19 infodemic. According to the World Health Organization, an infodemic is "too much information including false or misleading information in digital and physical environments during a disease outbreak" [6]. The COVID-19 infodemic has significantly contributed to vaccine hesitancy, which is the refusal of vaccines when access is not a limiting factor [7], across the United States. In the case of COVID-19, this hesitancy occurs despite an excellent vaccine safety profile. COVID-19 vaccine hesitancy is higher in some demographic groups who have even lower vaccination rates than the general population, demonstrating that the infodemic may disproportionately affect some communities [8,9].

Table 1. List of key terms and definitions.

Key terms	Definitions
Misinformation	False or incorrect information that is spread without malice
Disinformation	Inaccurate information that is spread deliberately with a deceitful or harmful intent
Infodemic	Excess amount of information on a topic that usually spreads rapidly and is confusing or unreliable
Vaccine confidence	Belief that vaccines are effective and safe
Vaccine hesitancy	Delay in acceptance or refusal of vaccinations despite vaccine availability
Vaccine refusal	Refusal of all vaccines including childhood vaccines

Social media platforms accelerated the dissemination of inaccurate information about the pandemic, contributing greatly to the COVID-19 infodemic and its health effects [10]. Social media has been shown to be more effective at promoting vaccine hesitancy than uptake, leading to a reduced effectiveness of public health measures and to decreased public engagement in disease prevention activities [7]. This was made worse by influential political representatives and cultural figures who spread misinformation and disinformation across all major social media platforms, which is referred to as top-down misinformation [11]. Although the connection between vaccination rates and the presence of misinformation online is known, best practices for mitigating the COVID-19 infodemic are yet to be determined. Thus far, there are no uniform efforts and policies by social media companies to combat harmful misinformation and disinformation present on their platforms. Additionally, the effect of a coordinated effort by social media companies to act against the COVID-19 infodemic is untested.

In August 2021, we hosted *INFODEMIC: A Stanford Conference on Social Media and COVID-19 Misinformation* to address issues related to the ongoing COVID-19 infodemic. Our primary aim for the conference was to develop best practices for social media companies to mitigate the COVID-19 infodemic online. In this paper, we report (1) the methods used to execute the virtual conference, (2) conference attendee engagement and analytics, (3) our qualitative analysis of the conference presentations, and (4) best practice recommendations.

Methods

Study Design, Setting, and Population

We conducted a thematic analysis of the transcript of a 6-hour, virtual conference about the COVID-19 infodemic. Our study aim was to identify best practices for social media companies to combat COVID-19 misinformation online; results of our thematic analysis represent the primary outcomes of the study. The transcript of the conference presentations formed the data for this study.

The conference was called *INFODEMIC: A Stanford Conference on Social Media and COVID-19 Misinformation*, and it was sponsored by Stanford University (Stanford, CA) [12]. It occurred on August 26, 2021, and data analysis was completed between October 2021 and November 2021. The study participants each consented to have their presentations recorded. These videos are now in the public domain; thus, no further consent was sought for this analysis. The Institutional Review Board (IRB) of Stanford University (IRB# 63151) deemed this study exempt.

The complete video recording of "INFODEMIC: A Stanford Conference on Social Media and COVID-19 Misinformation" is provided in Multimedia Appendix 1. This represents the raw data used for our qualitative thematic analysis.

Study Intervention

To meet our study aim, deliberate design and optimal execution of the conference were pivotal to the collection of meaningful data. We (RAB, MAG) convened 2 committees to assist with the project, one advisory and one for conference planning. First, we recruited a steering committee to oversee the project and provide strategic direction. The steering committee consisted of individuals from the following entities at Stanford University: the School of Medicine, the Internet Observatory, the Social Media Lab, the Digital Civil Society Lab, and the Health Communication Initiative. Members of the steering committee helped select and recruit some of the conference speakers, chose a date for the conference, and addressed day-of-conference logistics. Additionally, we recruited a planning committee (DC, TML, AR, AS, AX) that met weekly in the months leading up to the conference. The planning committee was comprised of Stanford students who successfully completed a course on the use of social media for knowledge translation in medicine. Students planned the conference under the direction of the steering committee and principal investigators.

We organized a conference agenda consisting of panel discussions and brief didactic presentations on a broad range of topics related to the COVID-19 infodemic and our study aim. We recruited expert speakers based on these topics from a variety of geographical, cultural, and academic backgrounds with special consideration of individuals who have power to influence vaccine-hesitant individuals. We used a diversity and inclusion lens to ensure a racial, ethnic, and gender balance of the participants in an attempt to mirror the general population who use social media platforms.

We created a conference website using Squarespace (Squarespace Inc, New York, NY) that served as the conference program, registration portal, and a platform to host 2 public calls for scholarly works [12]. The website and conference information were open-access, and conference registration was free. In addition to registration and advertising, the website promoted a call for research abstracts to be presented at the conference and a call for papers for a special theme issue of the *Journal of Medical Internet Research (JMIR)*: "Social Media, Ethics, and COVID-19 Misinformation." We scored submitted abstracts using Google Form (Google, Mountain View, CA) and selected the best submissions for presentation during a research symposium that immediately followed the conference. The research symposium served as a means of crowdsourcing additional information that might inform our study aim and allowed for presentation of relevant new research in the field. JMIR guest editors reviewed manuscript submissions to the theme issue separate from our study team.

We advertised the conference to potential attendees through social media and email outreach. We created Twitter (Twitter, San Francisco, CA) and Instagram (Meta Platforms Inc, Menlo Park, CA) accounts 4 months prior to the conference. We posted educational content, speaker spotlights, and promotional information to these accounts biweekly, increasing in frequency as the conference date approached. These accounts were also used for same-day conference backchanneling. We encouraged the participants to publicize the conference on their social media accounts, and we promoted the conference through Stanford Healthcare email distribution lists and newsletters.

Given COVID-19 travel restrictions, we held the conference virtually and opened registration to the public. We selected Hopin (Hopin Ltd, London, England) as our virtual conference platform because it provided important features necessary for moderated panel discussions: multiple speakers on the screen at the same time, audience interaction through chat or microphone, polling, and question and answer functions. We selected August 26, 2021 for the conference date based on our steering committee's accurate prediction of a COVID-19 vaccine surplus in the United States and the peak of news media attention about vaccine hesitancy. This date also minimized speaker conflicts and allowed enough time to plan the conference. INFODEMIC occurred 4 days after the US Food and Drug Administration formally approved the Pfizer (Pfizer Inc, New York, NY) COVID-19 vaccine, at a time of great public discourse over COVID-19 vaccine hesitancy.

On the day of the conference, we (RAB, MAG) met in person to manage conference logistics such as speaker login, backstage preparation, and online platform management. Virtually, other planning committee members (DC, TML, AR, AS, AX) oversaw the audience chat, monitored the questions and answers, took notes on key proceedings, and shared updates from the conference on social media. Abstract presentations followed the main conference program, as did a Stanford University student panel discussion on vaccine hesitancy led by the planning committee.

Participant Sampling and Data Collection and Analysis

We used a combination of purposive and snowball sampling to identify a diverse participant cohort [13,14]. We first determined the topics of the participant panel discussions based on the study aim, desired flow of the conference agenda, and format (eg, number of sessions, topics, length). We then identified expert participants for each panel via internet searches, past publications, the organizations they represented, social media platform searches, word-of-mouth recommendations, and steering committee recommendations. Our a priori search criteria identified participants for 7 of the 10 scheduled presentations. The remaining 3 presentations had participants who were referred to us by these experts (these included panels of social media representatives, government or religious leaders, and ethicists). We digitally recorded the conference and used Zoom (Zoom Video Communications Inc, San Jose, CA) to transcribe the 10 individual presentations.

Using a constructivist paradigm, we performed a thematic analysis of these transcripts to identify best practice recommendations for addressing COVID-19 misinformation and disinformation online [15-17]. We followed the 6-step framework for thematic analysis described by Braun and Clarke [17]. We inductively coded the transcript data for the existence of concepts related to COVID-19 misinformation, not the frequency they appeared in the transcript [17,18]. We analyzed

the transcripts to the level of sentences, grouped these responses into loose categories or concepts that were not predefined, and ignored irrelevant words [19]. Six study team members agreed on this preliminary coding approach and crafted rules before independently coding the content. We then met frequently to discuss code generation and meaning. Two study investigators were assigned to code each of the 10 transcripts separately, and these pairings were different for each transcript. Our final codebook consisted of codes agreed upon between the rater pairs. Using a consensus approach, we then conducted a team-based thematic analysis of the codes in a series of discussions among all of our investigators [15-17].

We used descriptive statistics to report measures of conference participation, engagement, and other analytics.

Reflexivity

We acknowledge that the experiences and opinions of our study investigators may have influenced our data analysis in this constructivist paradigm. Our senior investigator (MAG) is an emergency physician and medical education researcher who teaches a course on social media in medicine called, "Does Social Media Make Better Physicians?" Our study team includes 6 Stanford University undergraduate students (RAB, DC, TML, AR, AS, AX) who each completed that social media course and a subsequent research course that facilitated this study. Because of their common experiences in these 2 courses, all of our study investigators may have shared a preconceived perspective and understanding of the COVID-19 infodemic; we explicitly examined this bias throughout the generation of codes and themes. The students were novices to qualitative research; however, the senior investigator has ample experience with the analytic approach used and oversaw each methodological step. All coding was done in pairs of undergraduate students to balance prior experiences and assumptions. We designated our senior investigator as a potential third reviewer for the few coding disagreements that required adjudication. We maintained a research diary during the study that recorded all coding sessions, group discussions, and decisions during analysis of the transcripts. We reflected on the experience of engaging in this study at each team meeting.

Results

Conference Participation, Engagement, and Analytics

Participants

We recruited 26 conference participants with diverse expertise and occupations (Table 2). The participants were assigned to speak individually or in panels based on their areas of expertise (Table 3). We selected 9 abstracts for presentation in the research symposium. The authors represented 8 different countries.

Table 2. Demographic characteristics of the conference participants (age, gender, and race were not collected from participants).

Occupation (n=26)	Results, n (%)[a]
Physician	10 (38)
Ethicist	4 (15)
Social media influencer	4 (15)
Social media company representative	3 (12)
Public health expert	3 (12)
Politician	1 (4)
Religious leader	1 (4)

[a]Column does not total 100% because some participants had 2 occupations.

Table 3. Summary of the INFODEMIC conference presentations and participants.

Presentation title	Speaker(s)	Affiliation(s)	Presentation summary
Welcome Address	Seema Yasmin, MD	Stanford University	As the COVID-19 pandemic continues through its second year, the need for a strong public health infrastructure remains critical to mitigating the spread of misinformation. Partnerships with public health institutions are more effective than dissociating from them.
COVID-19 Update	Yvonne Maldonado, MD	Stanford University	Five billion vaccines have been administered worldwide, with 350 million administered in the United States alone. However, reaching marginalized populations remains a challenge, and access through community organizations must be a priority.
Vaccine Confidence	Heidi Larson, PhD	London School of Hygiene & Tropical Health	Much attention has been given to the sheer amount of misinformation flooding the public, but more important is the dynamic, fast-moving nature of misinformation that undermines credible science. Trust, rumors, and receptivity have played a large role in the disparate impacts and levels of confidence in misinformation observed between different countries.
Vaccine Hesitancy	Agnes Binagwaho, MD, PhD, MA; Gloria Giraldo, PhD, MPH; Aida Habtezion, MD, MSc; Seema Yasmin, MD (moderator)	University of Global Health Equity (AB); Latino Health Access (GG); Pfizer Inc (AH); Stanford University (AH, SY)	More than vaccine hesitancy, issues about vaccine equity and access must be addressed, especially because global pandemic inequities have been so apparent. Trust is the best way to minimize vaccine hesitancy—trust in government officials, trust in physicians, and trust in the local community. This trust cannot be built overnight but must be built in a grassroots, day-to-day manner. Global vaccine access disparities are alarming because of mismanaged vaccine distribution in high-income countries, leaving low-income, at-risk nations vulnerable.
COVID-19 and Distrust of Healthcare	Italo Brown, MD, MPH	Stanford University	There are significant racial disparities in vaccine acceptance, especially among the Black and Latinx communities. Combating these disparities means "fighting the misinformation Olympics," which refers to some communities having access to adequate information versus those who do not have credible information and trusted messengers. The embedded distrust of the health care system is an important etiology of vaccine hesitancy, and it must be contextualized in the history of communities of color, such as the devaluation of Black lives. There needs to be more communication of accurate information to communities of color, especially since information can often change and be distorted as it spreads. There also needs to be focus on restorative justice and validating mistrust; activating trusted messengers to enact change within communities and combat misinformation is essential.
Achieving COVID-19 Vaccine Equity	Tom Bollyky, JD; William A. Haseltine, PhD (moderator); Lisa Menning; Danielle Pacia, MBE	Council on Foreign Relations (TB); ACCESS Health International (WAH); World Health Organization (LM); The Hastings Center (DP)	A global approach to COVID-19 vaccinations is critical to ensure vaccine equity around the world. Thus far, vaccines have not been distributed to countries at highest risk, but instead 10 of the wealthiest countries are overrepresented in vaccine doses received. This is due to hoarding within wealthier countries, a lack of global cooperation and distribution, and a vaccine shortage worldwide. Knowledge sharing and infrastructure sharing are part of an important, related global disparity. In order to combat these issues, there needs to be greater transparency, more international cooperation, and delay of vaccine boosters until there is more universal coverage.

Presentation title	Speaker(s)	Affiliation(s)	Presentation summary
The Role of Social Media Companies	Aaron Berman; Brian Clarke; Renee DiResta (moderator); Anne Merritt, MD	Facebook (AB); Twitter (BC); Stanford Internet Observatory (RD); Google Search (AM); Stanford University (AM)	Twitter, Facebook, and Google each have policies for removing misleading content, adding warning labels, and deactivating accounts that promote misinformation. A major challenge for these companies is the removal of a huge influx of disinformation while promoting quickly evolving, high-quality, and medically accurate content. Collaborative efforts between social media companies and public health experts help fact-check and spread accurate information online. In the process of regulating online content, it is important for social media companies to balance free speech, censorship, and safety.
Leveraging Physician Influencers: The New Public Health Educators	Vin Gupta, MD, MPA	University of Washington; MSNBC Contributor	Social media platforms can be used effectively to tell stories and powerfully amplify accurate information, especially through visual media like videos and graphics. These platforms generate useful ideas, distribute actionable advice, build trust, and hold sources of misinformation accountable. Impactful, accurate medical information can be communicated online by leveraging multiple forms of media to engage different audience constituencies.
Do Social Media Influencers Affect Vaccination Rates?	Sanjay Juneja, MD; Jessica Malaty Rivera, MS; Cedric "Jamie" Rutland, MD	Baton Rouge General (SJ); Infectious Disease Expert at the COVID Tracking Project (JMR); West Lung (CR); Vice President of Association for Healthcare Social Media (CR)	Social media influencers can affect vaccination rates through effective science communication. Their messaging can make people confident about the vaccine, allow them to make choices out of a place of knowledge, and teach them to be discerners of truth. Empathy, respect, and relatability are crucial to empowering listeners to seek information, rather than antagonizing and shaming. Mitigating misinformation can come from debunking myths and anticipating logical fallacy traps people may espouse.
Role of Government and Religious Leaders in Mitigation Disinformation	Adrian Perkins; Gabriel Salguero, PhD; Matthew Strehlow, MD (moderator)	Mayor of Shreveport, Louisiana (AP); Reverend of the Gathering Place (GS); President of the National Latino Evangelical Coalition (GS); Stanford University (MS)	Combating COVID-19 disinformation relies heavily on collaborative partnerships. The partnership between public health and religious figures inspires trust and increases receptivity for the public, especially as trust in government officials declines. The politicization of the pandemic has contributed to distrust between opposing parties and threatened the mission of the faith community. Thus, spreading verified information and amplifying trusted voices within religious spaces are crucial to ensure that messages are communicated safely and reliably.
Ethical Imperatives for Social Media Companies and Influencers to Act	Nancy Berlinger, PhD (moderator); Arthur Caplan, PhD; Travis Rieder, PhD	The Hastings Center (NB); New York University (AC); Johns Hopkins University (TR); Berman Institute of Bioethics (TR)	Both effective science communication and ethical communication must be prioritized throughout the pandemic. The ethical imperatives of social media companies should include stricter self-regulation and refined presentations of information for the public. Social media companies are information disseminators, not journalists or reputable news people.

Same-Day Engagement

A total of 1090 conference registrants came from 71 different countries, and 587 registrants (587/1090, 53.9%) viewed the meeting live. Most registrations occurred during the days leading up to the conference (Figure 1). Peak attendance was 312 viewers (312/587, 53.2%) at any one time, and the average viewing time was 2 hours 44 minutes (164/360 contiguous minutes, 45.6%; Figure 2). The mean rating by attendees of the conference quality was 8.9 (lowest: 0 to highest: 10). Table 4 summarizes the complete conference analytics.

Figure 1. Most registrations occurred in the same week as the conference.

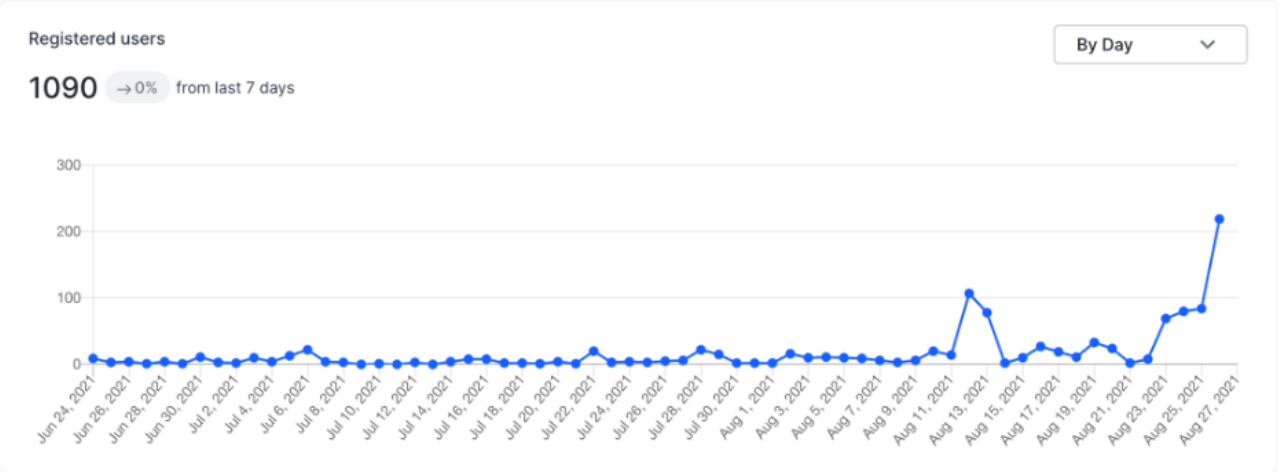


Figure 2. Peak active conference attendees by time.

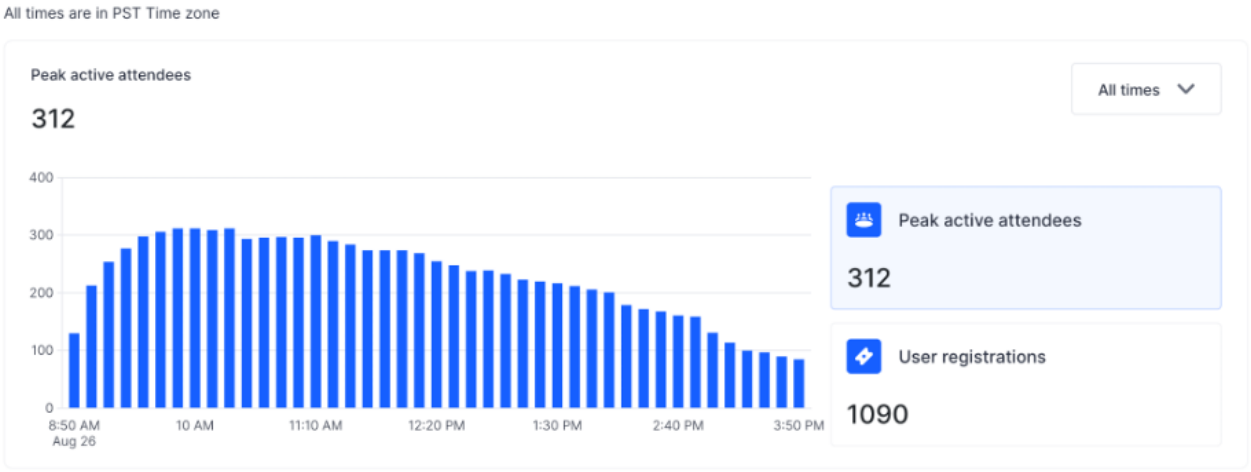


Table 4. Conference analytics as of December 1, 2021.

Analytics	Results
Registration analytics	
Registration fee	Free
Total registrations, n	1090
Final week registrations, n	486
Same-day registrations, n	219
International registrations, n	340
Countries with registrations, n	71
Top countries by registrations (n=1090), n (%)	
United States	747 (68.53)
Canada	39 (3.57)
Turkey	35 (3.21)
Philippines	23 (2.11)
United Kingdom	18 (1.65)
Other	228 (20.93)
Attendance analytics	
Total live attendees, n (% turnout)	587 (53.9)
Average attendee evaluation (out of 10; 10=high)	8.9
Peak attendance, n	312
Conference length (hours)	6
Average viewing time (minutes)	164
Asynchronous engagement, n	
Total YouTube viewers	929
Total Podbean listeners	9067
Instagram [20] followers	338
Twitter [21] followers	183
Unique website [12] views	5905
Conference links	
YouTube	[22]
Podbean	[23]

Asynchronous Engagement

We distributed a complete recording of the conference to all registrants for asynchronous viewing, including the 503 registrants who did not view the live broadcast. Additionally, a recording of the conference and its individual presentations is available via YouTube, and each presentation was recorded as individual podcast episodes hosted by the *Academic Life in Emergency Medicine* Podcast via Podbean (New York, NY) [22-24]. This adds asynchronous INFODEMIC "attendees" totaling 929 viewers and 9067 listeners as of December 1, 2021. Our conference Twitter and Instagram accounts had a combined following of 521 users and reached tens of thousands of others (Figure 3). Our website had approximately 5900 unique visitors and 12,000 page views (Figure 4).

Figure 3. Engagement with our social media accounts by date: (A) unique Instagram accounts reached and (B) Twitter impressions.

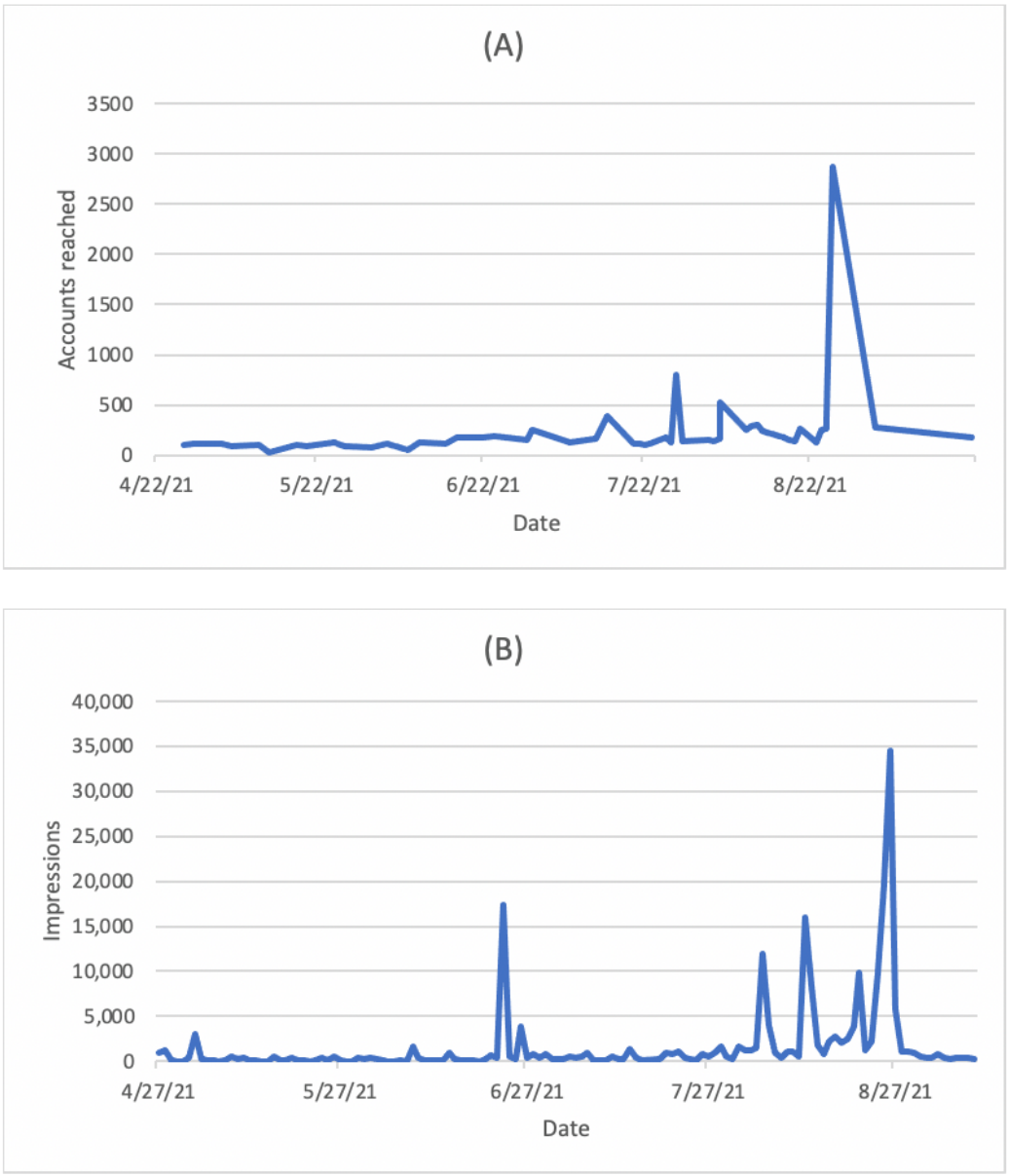
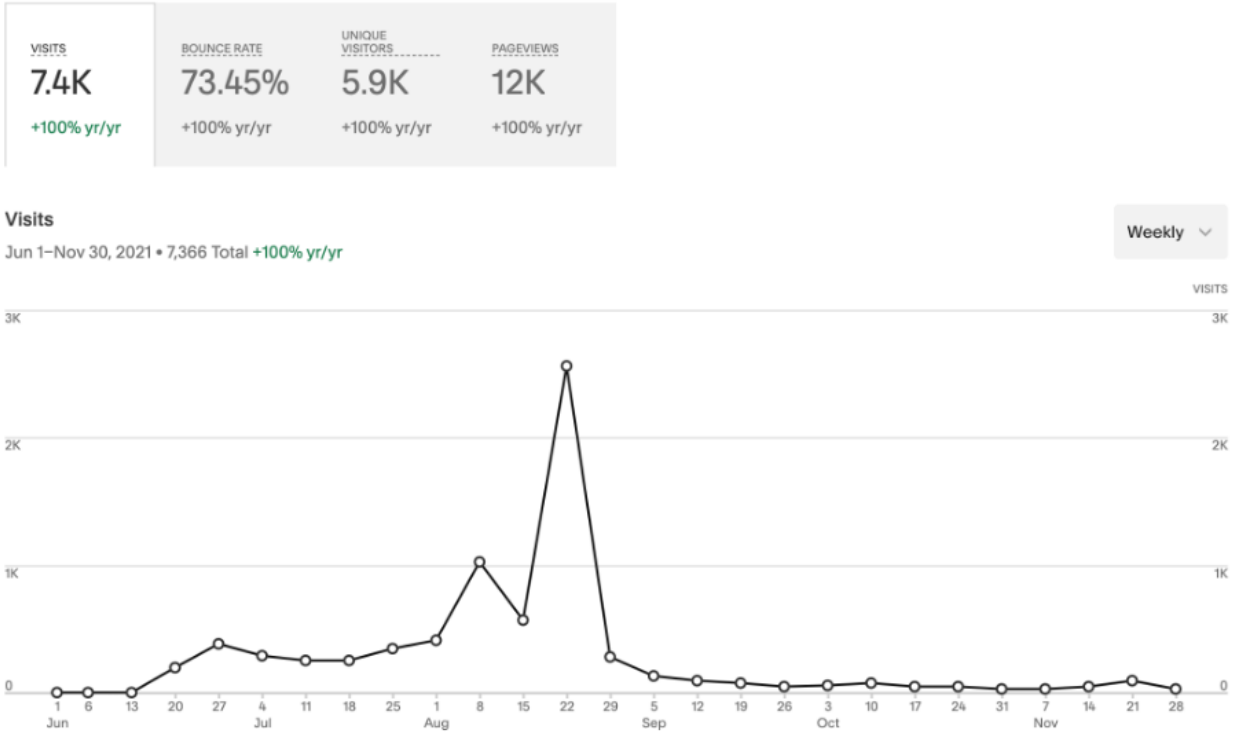


Figure 4. Website analytics showing a peak in visits during the week of the conference, August 22, 2021.



Thematic Analysis

Overview

The digital recordings of the 10 individual presentations resulted in 297 pages of transcribed text. Our analysis of this complete transcript resulted in 18 loosely associated categories of participant responses. These led to the identification of 4 major themes: trust, equity, social media practices, and interorganizational partnerships (Figure 5). These categories and themes broadly reflected the lack of access to both information and vaccines within marginalized populations, the global impact of COVID-19, and actions to minimize the spread of online misinformation (Table 5). Participants affirmed the ethical imperative on social media companies to use their platforms to curb the spread of COVID-19 misinformation because it continues to drive vaccine hesitancy and inequity.

Figure 5. Overview of our data analysis.

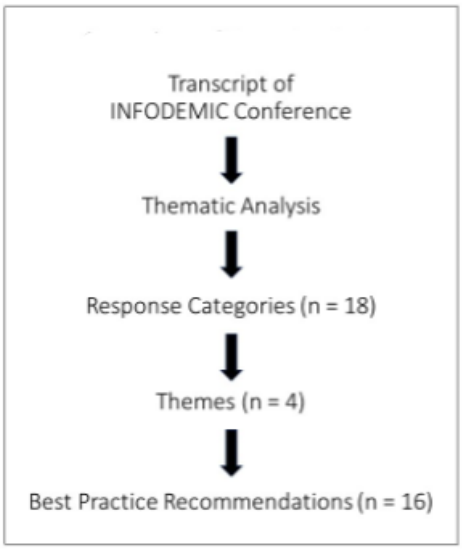


Table 5. The 4 themes and 18 response categories identified in the data analysis.

Theme	Categories
Trust	1. Historical mistreatment of marginalized populations contributes to their present-day distrust of health care. 2. Rumors, infodemics, and mistrust contribute to the spread of misinformation. 3. Community-based outreach is important for building trust. 4. Individual, collective, and commercial accountability is needed to reduce COVID-19 misinformation.
Equity	1. Inequitable access to accurate COVID-19 information fuels vaccine hesitancy and the disparate vaccination rates observed in some communities. 2. Inequitable access to accurate information results from the lack of representation of persons of color in communications and media. 3. COVID-19 vaccine access is more important than vaccine hesitancy globally. 4. Health disparities are directly related to social determinants of health.
Social media practices	1. Strategic design of social media platforms can reduce misinformation. 2. Empathy and respect are critical when communicating online with people who disagree with you. 3. Social media involves managing the different values and interests of users. 4. A balanced approach to free speech, censorship, and psychological safety is needed. 5. Social media companies can educate users by flagging and removing misinformation and promoting correct information. 6. Social media users can be taught how to critically analyze online information.
Interorganizational partnerships	1. Public health institutions should partner to combat the global impact of COVID-19. 2. Community health initiatives can fight COVID-19 misinformation. 3. Partnerships between social media companies and public health organizations can disseminate credible health information and fact-check online information. 4. Social media platforms can connect the science community with governmental agencies, commercial brands, and local health organizations.

Trust

Trust was the fundamental issue in discussions of COVID-19 vaccine hesitancy throughout the conference. Our participants described how disparities in COVID-19 vaccine acceptance disproportionately affected marginalized populations due to the "historical lack of trust by some communities because they have been mistreated by science, mistreated by politicians, mistreated by the judiciary system, and are still mistreated by [law enforcement]...so they don't trust." One participant recounted numerous instances where members of the Black community were mistreated, emphasizing how "the abuse of Black bodies creates the bedrock for this [modern] distrust" and "[currently] we still see elements of mistreatment in the medical system. So, it would make complete sense to know that Black communities have a lasting and lingering distrust of health care." Participants described how marginalized communities also suffer from a lack of "access to adequate information" because "having trusted messengers, people who can interpret messaging [and] can deliver adequate truthful information, is a commodity, [but] it is a rarity and it's scarce in [marginalized] communities."

One participant noted that "credible information and trusted voices have a viral nature, which [is] important in the context of the rampant mis- and disinformation" and that "a crucial factor that determines the spread of information is trust." The influence of rumors that threaten trust within different countries was emphasized, as in "one country, [rumors] fizzle out and die, [while] in another country, it disrupts a program and [shuts] down a vaccination initiative. It all depends on the receptivity and fertile ground [within each country that] either amplifies or mitigates the spread." Participants described "the process of building trust" through community-based organizations, which may not be an issue of "restoring trust but building trust with the public for the first time given long histories of exploitation." Several participants underscored the importance of ensuring individual, collective, and commercial accountability when dealing with COVID-19 misinformation on social media.

Equity

Our participants described how the COVID-19 pandemic highlighted countless global inequities in health care infrastructure and access to information and care. One participant said, "nothing to do with this pandemic is equal, from mortality and morbidity rate[s], from who's most susceptible to infection, [to] the inequitable access to vaccine." Participants said that social determinants of health such as poverty, frontline jobs, crowded living spaces, crowded public transit, food deserts, and inadequate insurance coverage can "lead to higher risk for COVID disease but also serve to alienate and isolate people from good information and leave them vulnerable to local pockets of misinformation." Another participant offered the example that, for pediatric deaths from COVID-19 around the world, "low- and middle-income countries have two and a half times the risk of deaths than high-income countries." The participants discussed how the pandemic has revealed this "uneven playing field," especially in denying equitable access to the vaccine among poorer countries. A participant cited a study done by the Africa CDC showing that 80% of people in African countries, 65% in the United States, and 60% in France are willing to get vaccinated, but "as of May 2021...only 2% of the African population has got access to the vaccine." Our participants agreed that vaccine access is more important than vaccine hesitancy in much of the world. These inequities not only are global but also plague

people within small communities, as one participant noted, "[sometimes] it's only a 15-mile difference" between having and not having access to a continuous medical system.

In the same way that inequities cause access challenges to vaccines and health care services, one participant explained that there are also some communities that are "more vulnerable to being targeted by disinformation campaigns and are more susceptible to believing misinformation." Our participants discussed how Pfizer focused on equitable access to vaccine information and clinical trials while developing their vaccine, describing that Pfizer put in the effort to "involve and ensure diverse populations are included." In addition to including minority groups in clinical trials, one participant described how the Pfizer vaccine development process involved partnering with "trusted civil societies and NGO [nongovernmental organizations] like the NAACP [National Association for the Advancement of Colored People], the National Black Nurses Association, the National Hispanic Medical Association" to approach vaccine development in an equitable and "culturally sensitive way," considering factors such as "removing barriers in enrollment like language."

Social Media and Best Practices

Social media companies and influencers both play important roles in limiting misinformation online, as well as spreading accurate and credible content. Participants acknowledged that online platforms including Twitter, Facebook, and Google all have policies for removing inaccurate information, flagging misleading content, and deactivating accounts that promote misinformation. For example, one participant shared that Twitter "remove[s] content with the highest propensity of harm and content that may invoke deliberate conspiracy theories relating to COVID-19 and COVID-19 vaccines," mitigating its spread. Social media companies have also been able to promote and amplify high-quality, medically accurate information via COVID-19 information pages and links labeling relevant posts. Through these efforts, Facebook has been able to "direct more than 2 billion people worldwide to expert health resources through [their] virus information center," and Google developed search features to "provide direct access to information from health authorities." Participants said that artificial intelligence systems can help "scale the impact" of misinformation and send it to "fact-checking partners around the world [that can] rate [the] content and reduce the distribution of it" if deemed harmful. Continual monitoring and tracking of online discussions allow companies to "have a better sense from a policy perspective where to start shifting [their] resources and where to start really understanding the patterns associated with it." The companies are working to navigate the challenge of regulating their platforms by finding a balance between "enabling people to express themselves freely while protecting the safety of [their] user community."

Health care influencers on social media can also impact vaccination rates by communicating information responsibly and effectively on their platforms. One way is by storytelling and giving actionable advice through a variety of media, whether that be "threads on Twitter, videos, [or] graphics"—all powerful ways of "leveraging the impact of multiple forms of media to their greatest extent" and increasing audience engagement. Influencers can use their platforms to build trust and generate ideas as ways to "transform people's minds, to bring them out of fear and into confidence about the vaccine—everything from flu to even COVID-19 and HPV [human papillomavirus]." Participants agreed that influencers can also help increase science literacy and teach viewers to make choices "out of a place of knowledge" and "increase people's ability to discern what is good data, what is bad data, and even how to read charts." Although educating viewers, it is crucial for health influencers to express empathy as an "undertone of the way in which [they] communicate...and empower people to seek this information out." Antagonizing and "shame-based motivation to seek information sends [people] to the darkest places of the internet" and is not as effective as coming from a place of understanding and respect.

Interorganizational Partnerships

The final theme that emerged from our conference was the dire need for interorganizational partnerships. From the first panel to the last, our participants emphasized the importance of partnering with like-minded organizations to improve public health and the information disseminated on social media. One participant advocated that health experts "should partner with public health and not dissociate ourselves from them because we need to build [the] public health infrastructure...[to] not only get through the pandemic, but all public health issues." They asserted the need to utilize "existing community engagement structures, local community leaders, faith-based organizations, community groups, and health centers" to "disseminate transparent and comprehensive information" for public good. These sentiments were echoed throughout the conference. As one participant stated, "partnerships are essential [and National Latino Evangelical Coalition] partners with the CDC and the Office of Health and Human Services to have webinars in the target language" to eliminate distrust in the health care system among the Latino community.

Lastly, participants agreed that social media companies have a responsibility to partner with health organizations to ensure that their platforms are "[addressing] a wide range of misinformation and the associated harms...[while] amplifying authoritative health information." One participant noted that Facebook is "partnering with other organizations to reach low vaccination rate communities such as campaigns featuring Black doctors or nurses or Spanish language campaigns" and "using that data that [they're] collecting in partnerships with academic institutions to judge the impact [on the pandemic] over time." Participants also discussed partnerships that Facebook maintains with health experts to "remove content that has been debunked as false and leading to physical harm" as well as partnerships with "the global network of more than 80 fact-checking organizations around the world in more than 60 languages" to identify misinformation about COVID-19 or the vaccine. Similarly, it was noted that Google has similar partnerships with "national public health authorities" to "display their content front and center" so users have "direct access" to it.

Discussion

Principal Findings

Health misinformation on social media poses a threat to public safety, especially in the midst of the COVID-19 pandemic. Our participants described how the increased use of social media during the pandemic paved the way for misinformation, conspiracy theories, and rumors to flourish, thus increasing fear and reluctance around the vaccine. With the emergence of new COVID-19 variants, it is especially vital that we combat vaccine hesitancy and vaccine inequity around the world. In particular, many of our participants acknowledged that social media companies have an ethical obligation to act against the spread of harmful misinformation on their platforms. Presently, social media companies are making efforts to mitigate misinformation—however, more could be done.

After analyzing transcripts from the INFODEMIC conference, we recommend 9 ways in which social media companies can better combat COVID-19 misinformation online (Table 6). While our intended goal was to make recommendations about COVID-19 misinformation to social media companies alone, our data were so rich that our analysis led to recommendations for health care professionals and the general public as well. During our analysis, we also identified several response categories specific to health care professionals and the general public, so we generated additional recommendations for each of these respective groups. Our findings are especially timely and important, as the literature has relatively few evidence-based guidelines for how social media companies should effectively manage the COVID-19 infodemic [25-27]. In addition, many health care professionals and social media users are unaware of the potential benefits of social media for the dissemination of factual scientific information and the promotion of COVID-19 vaccines [25,28,29].

Table 6. List of recommendations for social media companies, health care professionals, and the general public to mitigate COVID-19 misinformation.

Target	Recommendations
Social media companies	1. Increase representation by people of color as messengers of factual information to build trust in medicine and science within marginalized communities. 2. Promote posts from trusted sources of credible information. 3. Fact-check, flag, and remove posts that propagate the spread of misinformation. 4. Adjust search engines and other algorithms that push misinformation. 5. Use easily understandable information, infographics, and messaging. 6. Remind users to be critical of information they read online. 7. Partner with public health organizations to spread credible information. 8. Partner with one another in a coordinated effort to combat COVID-19 misinformation. 9. Encourage vaccinations and equitable access to vaccines.
Health care professionals	1. Engage in public health education online. 2. Use social media platforms to connect scientists, government officials, and health care organizations with the general public. 3. Be mindful of global health disparities when messaging about the vaccines.
General public	1. Critically analyze social media content related to COVID-19. 2. Disseminate scientific facts and evidence-based information. 3. Search for information rather than rely on social media algorithms to push content to you. 4. Be patient and empathetic in conversations with vaccine-hesitant individuals.

Some pre-COVID literature addressed health misinformation on social media platforms. Although these papers predate the current COVID-19 infodemic, our data support their overarching themes and recommendations. Chou et al [29] recommended a strong clinician-patient relationship for optimal health communication, and our findings underscore this importance of building patient trust in science, medicine, and health information messengers. Similarly, we found that individual and commercial accountability is needed to reduce COVID-19 misinformation. Chou et al [29] also recommended that social media companies implement mechanisms to validate information on their platforms. Our analysis yielded more detailed recommendations for social media companies to (1) flag and remove posts that propagate the spread of misinformation, (2) promote posts from trusted sources such as easily understood information and infographics, and (3) promote reminders to the public to be critical of what they see online [29]. Finally, Walter et al [25] performed a meta-analysis of pre-COVID studies that found that the most effective messenger of health information is a recognized expert in a given field, which aligns with the findings in our study as well.

Several unique aspects of the INFODEMIC conference planning and this subsequent study are noteworthy. The INFODEMIC planning committee and this author team were comprised primarily of undergraduate students (DC, TML, AR, AS, AX) who executed the event and this analysis as class projects. The students learned how to effectively work as a team, communicate with professional speakers, plan the logistics of a conference, operate the digital conference platform, and perform a qualitative analysis of conference transcripts. Second, our speaker selection was carefully planned so that a diverse group of fields was represented. For instance, we thought it was important to include religious leaders and physicians in discussions of social media regulations, a space normally dominated by social media companies and the government. The broad diversity of expert speakers was intentional, and this qualitative analysis was meant to discover commonalities among

their approaches to health misinformation. Finally, we believe that our thematic analysis of a full conference transcript is uncommon in the literature and represents a rigorous method of summarizing knowledge generated at a scientific meeting.

Lastly, our experiences hosting an online conference are instructive. The virtual platform allowed us to convene a panel of experts from locations in North America, Africa, and Europe simultaneously. It is unlikely that we could have recruited the same panel to a live conference due to costs, travel during a pandemic, and schedule availability. Conference registrants similarly represented an international audience. We also purposefully recorded the conference and made it freely accessible online through YouTube videos and Podbean podcasts. This ensures that the conference has continued impact, as evidenced by the 10-fold increase in asynchronous, on-demand viewers to date compared with live conference attendees. Additionally, registrations peaked in the final 48 hours prior to the conference, an unanticipated finding that is anecdotally common for virtual conferences. This may have been due to simultaneously higher social media activity about the conference in the days prior, both from the INFODEMIC planning committee and the invited speakers. These experiences are consistent with expert guidelines for planning virtual meetings [30,31].

Limitations

We acknowledge several important limitations of our qualitative study design. Most notably, our sampling technique is subject to selection bias, and thus, our data were contingent on the specific participants or panel discussion topics chosen. However, it is unclear if similarly selected content experts would have offered substantively divergent perspectives. We tried to address this issue by creating a diverse conference agenda, recruiting participants with content expertise, and performing a thematic analysis across all 10 of the conference presentations to increase the variety of participants' responses in the data set. Two notably absent participant types were vaccine-hesitant individuals and "antivaccine" individuals. Though we initially intended to include a vaccine-hesitant participant, we eventually chose not to recruit any because of the difficulty of ensuring their psychological safety during the conference. Additionally, we were concerned that an "antivax" (vaccine-refusing rather than vaccine-hesitant) participant might have led the conference astray from its primary objective. Such individuals may have interacted with our participants during their presentations anyway, using the virtual conference platform chat function; this is not clearly reflected in our data set as the transcripts captured audio-only content, not chat room discussions.

Additionally, several biases reflective of participant comfort during the live, internationally broadcast conference may have affected the results; these include social desirability bias, acquiescence bias, and sponsor bias. Each of these response biases were partially mitigated by the use of open-ended questions and panel moderation by participants rather than sponsors. Similarly, status quo bias may have affected the responses of participants representing social media companies. Some responses may have been further biased by the use of a virtual conference platform to host the conference. Although this format allows for a global audience, technology itself can be a barrier for participants and alter their engagement and responses. Moreover, the conference was held in English without closed captioning or translation services, which may have affected attendee interaction with the participants via the chat function. Finally, our results may be subject to confirmation bias and to the experiential bias of the co-investigators who planned the conference and wanted to ensure its success. These were addressed by the use of paired investigators to code data and through direct supervision and instruction by the first author who has expertise in this methodology.

Conclusions

The Stanford INFODEMIC conference brought together experts from various fields and backgrounds to discuss the etiologies of vaccine hesitancy and the power of social media to increase COVID-19 vaccination rates. There were common themes to the participant remarks: improving trust in science and medicine, promoting equity in vaccine access and health care, identifying social media best practices, and creating interorganizational partnerships. As the future course of the pandemic continues to be uncertain, we offered specific recommendations that social media companies, health care professionals, and the general public can adopt to better mitigate online health misinformation.

Acknowledgments

Funding for the conference came from a Stanford University Ethics, Society, and Technology Hub Seed Grant from the McCoy Family Center for Ethics in Society, Stanford University Introductory Seminars, and the Stanford University Department of Emergency Medicine. We thank the following individuals for their various contributions to the INFODEMIC Conference: Jonathan Berek, Samantha Bradshaw, Susan Coppa, Renee DiResta, Jeffrey Hancock, Kristen Kayser, Ashley Neis, Matthew Strehlow, and Seema Yasmin.

Authors' Contributions

MAG, RAB, DC, TL, AR, AS, and AX contributed to the study design, data collection and analysis, and manuscript preparation.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Stanford Infodemic Conference Recording.
[DOC File , 23 KB - jmir_v24i2e35707_app1.doc]

References

1. COVID-19 Dashboard. Johns Hopkins Coronavirus Research Center. URL: https://coronavirus.jhu.edu/map.html [accessed 2021-12-01]
2. Average number of daily deaths in the United States among the leading causes of death from March 2020 to November 2021. Statista. 2021 Dec. URL: https://www.statista.com/statistics/1254560/leading-causes-of-death-in-the-us-average-number-daily/ [accessed 2022-01-30]
3. COVID Data Tracker. Centers for Disease Control and Prevention. URL: https://covid.cdc.gov/covid-data-tracker/#datatracker-home [accessed 2022-01-30]
4. Ortaliza J, Orgera K, Amin K, Cox C. COVID-19 preventable mortality and leading cause of death ranking. Peterson Center on Healthcare and Kaiser Family Foundation. 2021 Dec 10. URL: https://www.healthsystemtracker.org/brief/covid19-and-other-leading-causes-of-death-in-the-us/ [accessed 2022-01-30]
5. Healy M. CDC shifts pandemic goals away from reaching herd immunity. Los Angeles Times. 2021 Nov 12. URL: https://www.latimes.com/science/story/2021-11-12/cdc-shifts-pandemic-goals-away-from-reaching-herd-immunity [accessed 2022-01-30]
6. Infodemic. World Health Organization. URL: https://www.who.int/health-topics/infodemic [accessed 2022-01-30]
7. Puri N, Coomes EA, Haghbayan H, Gunaratne K. Social media and vaccine hesitancy: new updates for the era of COVID-19 and globalized infectious diseases. Hum Vaccin Immunother 2020 Nov 01;16(11):2586-2593 [FREE Full text] [doi: 10.1080/21645515.2020.1780846] [Medline: 32693678]
8. Razai MS, Osama T, McKechnie DGJ, Majeed A. Covid-19 vaccine hesitancy among ethnic minority groups. BMJ 2021 Feb 26;372:n513. [doi: 10.1136/bmj.n513] [Medline: 33637577]
9. Fernández-Penny FE, Jolkovsky EL, Shofer FS, Hemmert KC, Valiuddin H, Uspal JE, et al. COVID-19 vaccine hesitancy among patients in two urban emergency departments. Acad Emerg Med 2021 Oct;28(10):1100-1107 [FREE Full text] [doi: 10.1111/acem.14376] [Medline: 34403539]
10. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. J Med Internet Res 2021 Jan 20;23(1):e17187 [FREE Full text] [doi: 10.2196/17187] [Medline: 33470931]
11. Brennan JS, Simon F, Howard PN, Nielsen RK. Types, sources, and claims of COVID-19 misinformation. The Reuters Institute for the Study of Journalism. URL: https://reutersinstitute.politics.ox.ac.uk/types-sources-and-claims-covid-19-misinformation [accessed 2022-01-30]
12. INFODEMIC: A Stanford Conference on Social Media and COVID-19 Misinformation. Stanford INFODEMIC. URL: https://www.stanfordinfodemic.org/ [accessed 2022-01-30]
13. Thompson SK. Adaptive sampling in behavioral surveys. NIDA Res Monogr 1997;167:296-319. [Medline: 9243567]
14. Creswell JW, Báez JC. 30 Essential Skills for the Qualitative Researcher. Newbury Park, CA: SAGE Publications, Inc; 2015.
15. Kiger ME, Varpio L. Thematic analysis of qualitative data: AMEE Guide No. 131. Medical Teacher 2020 May 01;42(8):846-854. [doi: 10.1080/0142159x.2020.1755030]
16. Miles MB, Huberman AM, Saldaña J. Qualitative Data Analysis: A Methods Sourcebook, 4th edition. Newbury Park, CA: SAGE Publications, Inc; 2018.
17. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
18. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. Nurs Health Sci 2013 Sep 11;15(3):398-405. [doi: 10.1111/nhs.12048] [Medline: 23480423]
19. Kuckartz U. Qualitative Text Analysis: A Guide to Methods, Practice & Using Software. Newbury Park, CA: SAGE Publications Ltd; 2014.
20. Stanford INFODEMIC. Instagram. URL: https://www.instagram.com/stanfordinfodemic/ [accessed 2022-01-30]
21. Stanford INFODEMIC. Twitter. URL: https://twitter.com/suinfodemic?lang=en [accessed 2022-01-30]
22. INFODEMIC: A Stanford Conference on Social Media and COVID-19 Misinformation. YouTube. 2021 Sep 02. URL: https://youtu.be/gYTt8pRB5d0 [accessed 2022-01-30]
23. Gisondi M, Barber R. Stanford's INFODEMIC Conference on COVID-19 Misinformation: Open-access podcasts. ALiEM. URL: https://www.aliem.com/infodemic-conference-covid-19-misinformation-podcasts/ [accessed 2022-01-30]
24. Podbean. Academic Life in Emergency Medicine. URL: https://academiclifeinem.podbean.com/ [accessed 2022-01-30]
25. Walter N, Brooks JJ, Saucier CJ, Suresh S. Evaluating the impact of attempts to correct health misinformation on social media: a meta-analysis. Health Commun 2021 Nov;36(13):1776-1784. [doi: 10.1080/10410236.2020.1794553] [Medline: 32762260]
26. Gottlieb M, Dyer S. Information and disinformation: social media in the COVID-19 crisis. Acad Emerg Med 2020 Jul;27(7):640-641 [FREE Full text] [doi: 10.1111/acem.14036] [Medline: 32474977]

27. Sheng AY, Gottlieb M, Welsh L. Leveraging learner-centered educational frameworks to combat health mis/disinformation. AEM Educ Train 2021 Aug 04;5(4):e10711. [doi: 10.1002/aet2.10711] [Medline: 34926972]
28. Bautista JR, Zhang Y, Gwizdka J. Healthcare professionals' acts of correcting health misinformation on social media. Int J Med Inform 2021 Apr;148:104375. [doi: 10.1016/j.ijmedinf.2021.104375] [Medline: 33461008]
29. Chou WS, Oh A, Klein WMP. Addressing health-related misinformation on social media. JAMA 2018 Dec 18;320(23):2417-2418. [doi: 10.1001/jama.2018.16865] [Medline: 30428002]
30. Howard M. The complete guide to virtual events in 2021. Cvent. URL: https://www.cvent.com/en/blog/events/virtual-events [accessed 2022-01-30]
31. The Definitive Guide to Event Registration. Aventri. URL: https://www.aventri.com/strategy/event-registration-strategy [accessed 2022-01-30]

Abbreviations

CDC: Center for Disease Control and Prevention
HPV: human papillomavirus
IRB: institutional review board
JMIR: Journal of Medical Internet Research
NAACP: National Association for the Advancement of Colored People
NGO: nongovernmental organizations

Edited by A Raja; submitted 16.12.21; peer-reviewed by A Sheng, A Wahbeh, H Ayatollahi; comments to author 02.01.22; revised version received 09.01.22; accepted 09.01.22; published 15.02.22.

Please cite as:
Gisondi MA, Chambers D, La TM, Ryan A, Shankar A, Xue A, Barber RA
A Stanford Conference on Social Media, Ethics, and COVID-19 Misinformation (INFODEMIC): Qualitative Thematic Analysis
J Med Internet Res 2022;24(2):e35707
URL: https://www.jmir.org/2022/2/e35707
doi:10.2196/35707
PMID:35030089

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org